Competing for Inventors: Market Concentration and the Misallocation of Innovative Talent*

Andrea Manera†

January 13, 2022

LINK TO LATEST VERSION

Abstract

The rapid productivity gains achieved by technological innovations in the 20th century have slowed in recent decades. This has come at a time of increased market concentration. In this paper, I explore how dominant companies in concentrated sectors have siphoned off inventors that might have been employed more productively in competitive industries. For the period 1997-2012, I establish that sectors with rising concentration captured a disproportionate share of researchers, while also experiencing a decrease in R&D productivity, signaled by falling forward citations and slowing growth per inventor. These findings imply that inventors became increasingly misallocated, accounting for nearly 30 percent of the decline in the average annual growth rate of output per worker over the 15-year study period. I show that these results arise naturally in a Schumpeterian growth model where monopolistic firms conduct "defensive patenting" to hamper competitors' R&D. A calibration of this model reveals that a planner interested in maximizing growth should allocate R&D tax credits to entrants in high-concentration sectors.

JEL Codes: O30, O31, O32, O40.

Keywords: Market Concentration, Defensive Innovation, R&D Productivity.

*I am extremely grateful to my advisors Daron Acemoglu, Martin Beraja, and John Van Reenen for invaluable advice and support. I am also grateful to Michele Fornino for lending indispensable computing power and for his collaboration on a previous version of this project, and Nicholas Zolas, who kindly shared with me the ALP classifications of patents into NAICS sectors of application. I also thank Josh Angrist, David Autor, Pierre Azoulay, Ricardo Caballero, Arnaud Costinot, Joel Flynn, Germán Gutiérrez, Clémence Idoux, John Sturm, Martina Uccioli, Iván Werning, Christian Wolf, and participants in the MIT Macro Lunch, Applied Micro Lunch, and Macro Seminar for useful comments and suggestions.

†Massachusetts Institute of Technology, Department of Economics, E52-480, 50 Memorial Dr, Cambridge, MA 02142. Email: manera@mit.edu.

1 Introduction

Research and Development activities, which are key to innovation and growth, have generated increasingly smaller gains in productivity over the last few decades (Bloom et al., 2020; Fernald and Jones, 2014; Gordon, 2016). A prominent explanation: Technological complexity is making it harder to come up with new ideas. A less explored alternative is misallocation of R&D resources (Acemoglu et al., 2018, 2021), a concern that is echoed in policy circles and the press (Metz, 2017; Bass and Brustein, 2020). At issue, for example, is whether dominant high-tech firms are attracting a disproportionate share of highly educated, highly skilled workers at the expense of companies in more competitive sectors. According to TalentSeer (2020), 20% of total Artificial Intelligence experts are employed by just five big-tech companies: Google, Microsoft, Apple, Amazon, and IBM. At the same time, smaller firms in other sectors appear unable to attract this scarce talent. Two natural questions then arise: Is such allocation inefficient? And if so, can inventor misallocation explain the observed fall in R&D productivity?

To answer these questions, I study the broader effect of increasing market concentration on the allocation of inventors across sectors. I start by documenting several novel facts using US Patent and Trademark Office (USPTO) data and concentration measures from the Economic Census over the period 1997-2012. First and foremost, I show a positive correlation between increases in a sector's market concentration and the share of inventors it attracts. This correlation might indicate either that firms in high-concentration sectors have drawn more inventors or that they owe their success to large investment in research. To address this reverse causality problem, I adopt an instrumental variable strategy. I use the increase in the number of product-market regulations from Mercatus RegData as an instrument for increased concentration. This specification shows that sectors with increased concentration have increased their share of inventors and not vice versa.

Two additional findings substantiate that high-concentration sectors are using R&D resources inefficiently and consequently depressing aggregate research productivity. First, the quality of patents in sectors with increased concentration declined, as measured by patent forward citations. Second, research productivity, as measured as growth in output per worker per inventor, has decreased in these sectors.[1] Further, I find that dominant firms draw a disproportionate share of their sector's inventors. Based on these three observations, I conclude that incumbent firms are focusing their research efforts on "defensive innovation," that is, projects with the primary aim of warding off potential competition. Quantitatively, the midpoint of my estimates implies that inventor misallocation is responsible for a 0.78pp fall in the average annual growth in output-per-worker growth in the study sample. That

[1] This measure of inventor productivity mirrors the analogous definition of "research productivity" offered by Bloom et al. (2020). In the introduction to their paper, they decompose economic growth into the product of the number of researchers and a term capturing research productivity. Following this definition, I compute inventors' productivity as the ratio of the growth in each sector, which I measure using output per worker growth, and the number of inventors employed by that sector.

translates to 27.3% of the observed decline over the period 1997-2012.

Methodologically, my analysis relies on a novel dataset of "knowledge markets," defined as sets of product markets that share similar inventors. These markets are based on the network of inventor transitions across product categories, identified using USPTO patents classified by their NAICS sectors of application. This approach avoids pooling inventors with unrelated technical expertise, which would bias the response of inventor mobility to sectoral characteristics toward zero.

In the second part of the paper, I develop a model of R&D resource allocation. My objectives are twofold. First, I use the model to explain how inefficient defensive innovations can arise and proliferate as concentration increases. Second, I quantify the R&D productivity effects of increasing concentration and evaluate how R&D subsidies can be best allocated in the presence of defensive innovation. I adopt a Schumpeterian creative-destruction framework in which new entrants conduct productive R&D while incumbents employ inventors in defensive projects. A two-sector general equilibrium model shows that unbalanced changes in concentration across sectors generate a fall in inventor productivity and growth. Indeed, inventors shift to less competitive markets, where defensive projects, which hamper entry and Schumpeterian growth, are more prevalent, and away from competitive sectors where their efforts would be more productive. The theoretical analysis shows that defensive innovation is the key factor behind increased concentration of inventors among incumbents and the fall in R&D productivity.

I calibrate a two-sector version of my model to match moments of the R&D spending distribution in 1997 and growth over the period 1997-2012. This calibration produces a 2.5% fall in output per worker growth from misallocation, close to the 3% lower bound implied by my estimates. In the context of this model, I study the allocation of cost-neutral R&D subsidies that maximize growth. The model suggests that subsidizing entrants' R&D in more concentrated sectors constitutes the most effective policy, leading to a rise in annual GDP growth of about 0.5pp (a 17% increase from the 2012 benchmark). Similar results obtain if entry is uniformly subsidized across sectors. This finding resonates with the fact that defensive innovation is the main inefficiency in the model. Since this friction acts through an increase in entry barriers, the best way to counter it is to lower entry costs.

The rest of the paper proceeds as follows. In the following section, I survey the related literature place my study in its context. Section 2 describes my data sources, focusing in particular on the construction of knowledge markets. Section 3 reports the results of my empirical analysis. Section 4 details my theoretical framework, and conducts the policy analysis. Section 5 concludes and provides directions for future work.

1.1 Related Literature

My work builds on the empirical and theoretical literature analyzing the impact of product market competition on innovation. First, I analyze how increased concentration affects the allocation of R&D *across sectors* in a departure from previous work that took either a within-sector or economy-wide approach. Second, I explicitly analyze the allocation of R&D inputs across sectors by identifying the boundaries of markets for inventors rather than relying on R&D spending or similar measures of innovative effort as in previous work. Third, I correlate changes in the competitive environment to documented trends on falling R&D productivity. Fourth, I model the impact on R&D productivity of pre-emptive innovation in a multi-sector model, and I arrive at a growth-maximizing policy in this context. Both the focus on R&D productivity and the policy analysis are novel to the theoretical literature analyzing defensive innovation.

First an foremost, my paper relates to the vast literature estimating the empirical effect of competition on innovation. Aghion et al. (2005) found that innovation increased at low levels of competition and decreased at high levels, depicting the relationship as an inverted U. Accordingly, papers in this literature have highlighted contrasting effects of competition on overall R&D activity, focusing mostly on episodes of trade liberalization(see Shu and Steinwender, 2019, for an extensive review). Most papers in this strand identify these effects at the firm-level, which restricts their scope to the effect of competition within product markets. My paper instead adopts a cross-sector view, analyzing the extent to which decreased competition in one market draws away resources from other markets. To do so, I build a novel dataset of "knowledge markets," sets of product markets that share the same inventors. While several papers investigate the mobility of inventors (see, e.g., Azoulay et al., 2017; Moretti and Wilson, 2017), I believe mine to be the first to analyze the effects of market structure on inventors' movements across sectors.

With its focus on competition and innovation, my paper connects to literatures that document increased concentration (Autor et al., Forthcoming; Gutiérrez and Philippon, 2017; Grullon et al., 2019); profits and markups (Barkai, 2020; De Loecker et al., 2020; Eggertsson et al., 2018); and the relationship between falling innovation and R&D productivity (Akcigit and Ates, 2019, 2020, 2021; Bloom et al., 2020) and the allocation of R&D within and across sectors (Acemoglu et al., 2018, 2021; Akcigit and Kerr, 2018). My contribution bridges these literatures, explicitly linking changes in the competitive structure to the allocation of R&D resources across more and less concentrated sectors, and their deployment to productive or defensive projects.

Several papers document the role of pre-emptive innovation in ordinary firm operations (see Guellec et al., 2012, for a review of the evidence), and the high valuation of the resulting patents (Abrams et al., 2013; Czarnitzki et al., 2020; Grimpe and Hussinger, 2008). Most recently, Argente et al. (2020) show that, within product markets, large firms tend to account for the bulk of patenting activity, but are responsible for a smaller share of implemented innovations relative to non-patenting firms. The authors interpret this finding as evidence of defensive innovation, intended to deter competition. My paper builds on this literature showing that increased concentration raises the incentives for defensive innovation, as demonstrated by a fall in forward citations in concentrating sectors. This result connects to the findings and the model of Abrams et al. (2013), who study the cross-sectional relation between patent value and forward citations theoretically, showing that high-value patents also tend to receive fewer citations, and rationalize this result through pre-emptive innovation.

On a theoretical standpoint, I embed defensive patenting as modeled in in (Abrams et al., 2013) into a Schumpeterian growth model, building on the extensive literature inaugurated by Aghion and Howitt (1992). My solution relies on several results derived by Acemoglu and Akcigit (2012). To the best of my knowledge, my paper is the first to analyze the impact of defensive innovation in the context of a general-equilibrium growth model. The closest precedent to this analysis is Jo (2019), who builds on Akcigit and Kerr (2018), and characterizes incremental innovation, that is refinement of existing product lines, as "defensive." Therefore, in his framework, defensive innovation aims to increase the technological distance of incumbents from entrants, in the tradition of Aghion et al. (2001). By contrast, in my framework defensive innovation is specifically aimed at protecting dominant positions and reducing entry as in Abrams et al. (2013). I extend their model to consider the effects of defensive innovation on R&D productivity and overall innovation. My final contribution consists in analyzing the growth-maximizing allocation of R&D subsidies, which has not been previously studied in this context.

2 Data description

In this section, I detail source material and the construction of the dataset used in my empirical analysis. Subsection 2.1 lists the sources of the raw data. Subsection 2.2 focuses on the definition of inventor productivity measures and knowledge markets, which I identify through realized inventor flows across sectors. Subsection 2.3 briefly describes other data construction steps that are discussed in more detail in Appendix A.

2.1 Data Sources

My empirical analysis relies on the variation of concentration across product markets, as defined by 4-digit NAICS sectors, the impact of these shifts on the allocation of inventors with specific competences across these sectors, and the subsequent effect on inventor productivity. I use USPTO patent data to measure inventor productivity and establish the set of product markets that share similar inventors, and US Economic Census data to obtain concentration and productivity growth measures. Finally, I also use a dataset of product market regulations, Mercatus RegData 4.0, to conduct an instrumental variable analysis, as well as NBER-CES to obtain estimates of the Lerner Index that I employ in the calibration of my theoretical model.

My primary source is USPTO patent data from PatentsView. This dataset contains disambiguated patent, inventor, and assignee identifiers, as well as Cooperative Patent Classification (CPC) classes for each of the patents registered from1975 to 2021. I then identify inventor flows across different sectors, employing the ALP classification of 1976-2016 patents into NAICS sectors of application developed by Goldschlag et al. (2016). Since this classification is constructed using the PATSTAT dataset, I rely on the crosswalk built by Gianluca Tarasconi to match these two sources.[2] This leaves me with one third of all the patents registered between 1975 and 2021, due to the restriction of the time frame to 1976-2016 and an incomplete match between PATSTAT and PatentsView. I comb patent records for self-citations, truncation-corrected forward citations, and patent generality, following the procedure in Hall et al. (2001) and Acemoglu et al. (Forthcoming). I restrict my attention to utility patents, as I am interested in patents with a technological content and not just design improvements.

My main source for concentration and sales data is the US Economic Census (EC), which reports sales shares for the top 4, 8, 20, and 50 firms; the Herfindal-Hirschman Index; sales and number of companies in various NAICS 4-digit sectors at a 5-year frequency. I restrict my attention to the period between 1997 and 2012 for three main reasons. First, as I show below, this period saw substantial increase in the concentration of inventors in specific technology classes. Second, the start of this period coincides with an acceleration in the growth of market concentration and markups (see, e.g., De Loecker et al., 2020). Third, 1997 saw the adoption of the NAICS classification, thus ensuring a consistent definition of product markets throughout the period I analyze As my baseline concentration measure, I rely on the HHI lower bound constructed by Keil (2017).[3] I did so because the Economic Census reports the HHI only for a subset of industries, which would severely limit my sample. The method proposed by Keil (2017) obviates this issue by constructing the implied lower bound of the HHI implied by top sales shares reported in the Economic Census, which are available for a much wider set of industries than the HHI.[4] While my estimates are robust to using the EC-reported HHI, this choice allows me to obtain more power for my findings as well as to generalize them.

The Economic Census also provides sector-level growth in output per worker, which constitutes my main measure of productivity growth. I choose this measure instead of multi-factor productivity since the latter is available only for a limited set of sectors, mostly in manufacturing. I deflate sales using NAICS-specific price indices from the Bureau of Labor Statistics.

I employ two additional data sources in the empirical analysis and in the calibration of my model. First, I obtain sector-specific counts of regulations for various NAICS 4-digit sector from the Mercatus

[2] See https://patentsview.org/forum/7/topic/143, https://rawpatentdata.blogspot.com/2019/07/patstat-patentsview-concordance-update.html

[3] Available at https://sites.google.com/site/drjankeil/data.

[4] As detailed in Keil (2017) this measure is very strongly correlated with the HHI reported by the Economic Census when this is available, with a correlation of around 0.93.

RegData 4.0 dataset. I employ them to conduct an instrumental variable analysis, strengthening the causal interpretation of my results.[5] Second, I use NBER-CES data to produce estimates of the Lerner Index following Grullon et al. (2019).

All told, out of a total of 304 NAICS 4-digit sectors, I have assembled 157 business sectors for which I can measure the interrelation between concentration and knowledge markets.

2.2 Effective Inventors and Knowledge Markets

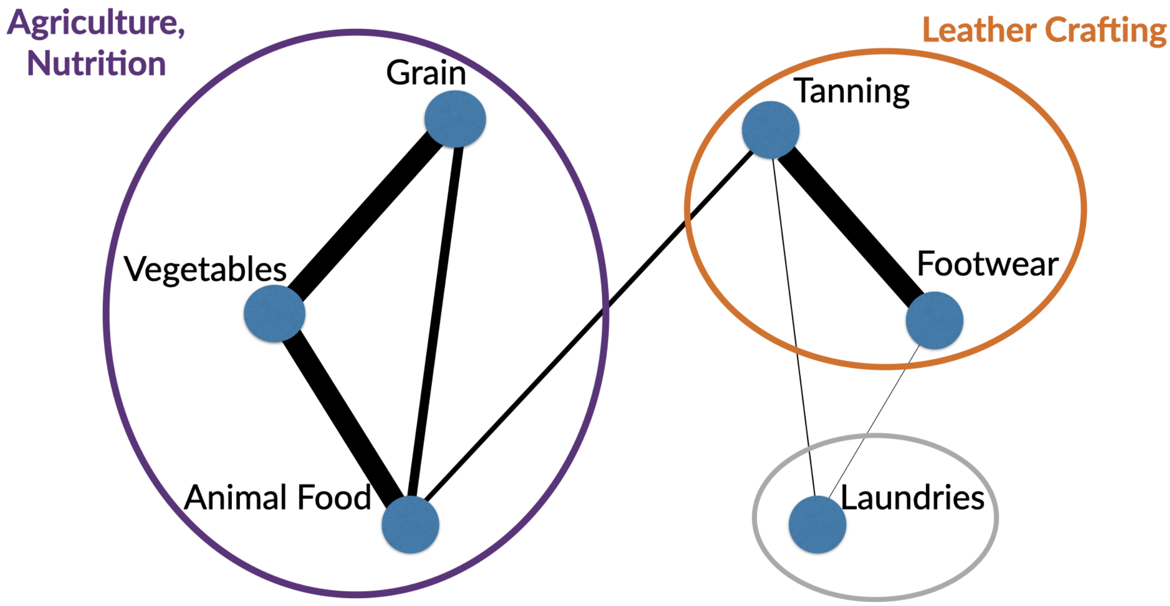
The main aim of this section is grouping product markets that share the same *required knowledge to innovate* and therefore compete for the same R&D inputs, namely inventors. I identify sectors that routinely exchange researchers through the Louvain community-detection algorithm (Blondel et al., 2008).

Figure 1 illustrates how I construct knowledge markets. Each node in the Figure represents a different NAICS 4-digit sector, and the black lines designate the inventor flows. The procedure shows how I determine these flows and measure their strength. After obtaining these weighted flows, I employ a community detection algorithm to group together sectors most closely connected. Figure 1 depicts strong flows among grain, vegetable farming and animal food manufacturing, all of which involve knowledge related to agriculture and nutrition, and separately between footwear and tanning, which both require knowledge of leather crafting. In this case, my algorithm would identify two knowledge markets, one given by the agriculture and food manufacturing sectors, and the other by leather crafting sectors, leaving the laundry services sector isolated. Based on the strength of connections, we would expect increased concentration in footwear manufacturing to attract inventors away from leather tanning, but not from vegetable farming. Increased concentration in the laundry sector, which has weak ties to the other sectors, would lead to negligible inventor movement within this particular grouping.

Measuring Inventor Transitions I employ the USPTO patent data classified into 4-digit NAICS sectors by Goldschlag et al. (2016) to construct knowledge markets. Table 1 depicts a hypothetical matching of the USPTO dataset with NAICS classifications. Note that the first patent has multiple inventors and is applicable to multiple sectors. Inventors are each assigned a disambiguated ID corresponding to the serial number of their first patent. In this example, inventor 00001-1 and 00001-2 both cooperate on the development of patent US00001. The third column in Table 1 shows the Goldschlag et al. (2016) classification for NAICS 4-digit industries. This classification is not limited to a single sector per patent, and includes multiple sectors in almost all instances. For instance, patent US00001 relates to multiple sectors, while patent US00002 is applicable to just one sector. Importantly, this classification captures

[5] Available at https://www.quantgov.org/bulk-download.

Figure 1: Graphical Illustration of Knowledge Markets



Note: This figure provides a graphical illustration of the definition of knowledge markets as sets of product markets sharing the same required knowledge to innovate. This illustration is based on transitions of inventors across product markets observed in my data and classified in the same knowledge markets, although many other sectors in these markets are excluded for the sake of exposition. In the figure, nodes represent NAICS sectors 1111 (Oilseed and Grain Farming), 1112 (Vegetable and Melon Farming), 3111 (Animal Food Manufacturing), 3161 (Leather and Hide Tanning and Finishing), 3162 (Footwear Manufacturing), and 8123 (Drycleaning and Laundry Services). The edges connecting nodes represent inventor transitions across sectors, while the width of these edges represents the strength of the connection between the two sectors as measured by undirected inventor flows.

the *technological nature* of the patent and the sectors of application of the knowledge required to develop that patent. While other classifications, like the CPC or the USPC, also describe the technological nature of patents, they do not allow a direct match to sectors of application.

Given this data structure, I define a transition in two ways. First, I consider inventor transitions *within patents.* That is, I consider that an inventor transition occurs between two sectors if an inventor works on a patent that applies to both. The direction of flows does not matter for the definition of knowledge markets, since I am only interested in grouping sectors that exchange researchers. Table 1 depicts two transitions between sectors 1111 and 1112 in 1980. The second type of transition that I consider is *across patents.* This transition occurs when an inventor applies his knowledge to patents in different product markets, such as between sector 1112 and 3111 by inventor 00001-1. The raw count of transitions of inventors across sectors in each year constitutes the basis of my measure of inventor flows.

Table 1: USPTO Data Structure

Patent ID	Inventor ID	Goldschlag et al. (2016) NAICS	Year
US00001	00001-1	1111	1980
US00001	00001-1	1112	1980
US00001	00001-2	1111	1980
US00001	00001-2	1112	1980
US00002	00001-1	3111	1981

Note: This table displays a hypothetical example of the data structure employed to build knowledge markets. The columns "Patent ID" and "Inventor ID" represent disambiguated patent and inventor identifiers as classified by USPTO PatentsView Data. The column "Goldschlag et al. (2016) NAICS" classifies patents into NAICS 4-digit sectors.

Weighting Inventor Flows: Effective Inventors After identifying transitions, I proceed to weigh them by two alternative measures in order to assess the flow of inventors across sectors. The first measure weighs each transition equally, computing inventor flows as the raw count of researchers moving across NAICS. The second measure adjusts for the productivity of individual inventors, since raw counts might overstate or understate the importance of each transition, depending on the size of origin and destination sectors, their technological nature, as well as the proficiency of each inventor. I therefore define a measure of "effective inventors" that aims to correct for these and other omitted factors. For each inventor,

I estimate the fixed effect, α_i, in the fully-saturated regression

$$\#\text{Patents}_{cfit} = \alpha_i + \gamma_{cft} + \varepsilon_{cfit}, \tag{1}$$

where $\#\text{Patents}_{cfit}$ denotes the number of patents registered in CPC class c; firm (assignee) f; and year t, that include inventor i. In this regression γ_{cft} denotes a of CPC class by firm (assignee) by year fixed effect. I choose to include indicators for one-digit CPC classes, the broadest classification, to identify as many fixed effects as possible. The fixed effect γ_{cft} controls for specific technological features of the patented technology, the firm environment, as well as the year. Further, this specification produces an estimate of inventor productivity that accounts for the number of collaborators on each patent. Given this specification, I define an *effective inventor* as one unit of the resulting fixed effect α_i, rescaled to take nonnegative values. Since these fixed effects might be inconsistently estimated, I check the robustness of all my results, including the construction of knowledge markets, to the use of the raw count of inventors.

Armed with the results of this estimate, I define *effective inventor flows* between sector j and sector k at time t as:

$$flow_{j\to k,t} = \sum_i \#\{i\text{'s transitions } j \to k \text{ in } t\} \cdot \alpha_i,$$

that is, the sum of transition counts weighted by effective inventors. The total undirected flow between two sectors is then given by the sum of inflows and outflows with ends in one of the two sectors:

$$flow_{jk} = \sum_t \left(flow_{j\to k,t} + flow_{k\to j,t} \right).$$

This flow measure defines a network of inventor transitions across product markets, where the nodes, j, k, are given by 4-digit NAICS codes, edges are given by transitions across sectors, and edge weights are defined as a rescaled version of $flow_{jk}$. I use these edge weights as a measure of the strength of the connection between pairs of sectors in the network. Rescaling the flow measure is necessary in order to exclude effects of sector size as well as to avoid double counting of inventors. I describe how I rescale this series in Appendix A.

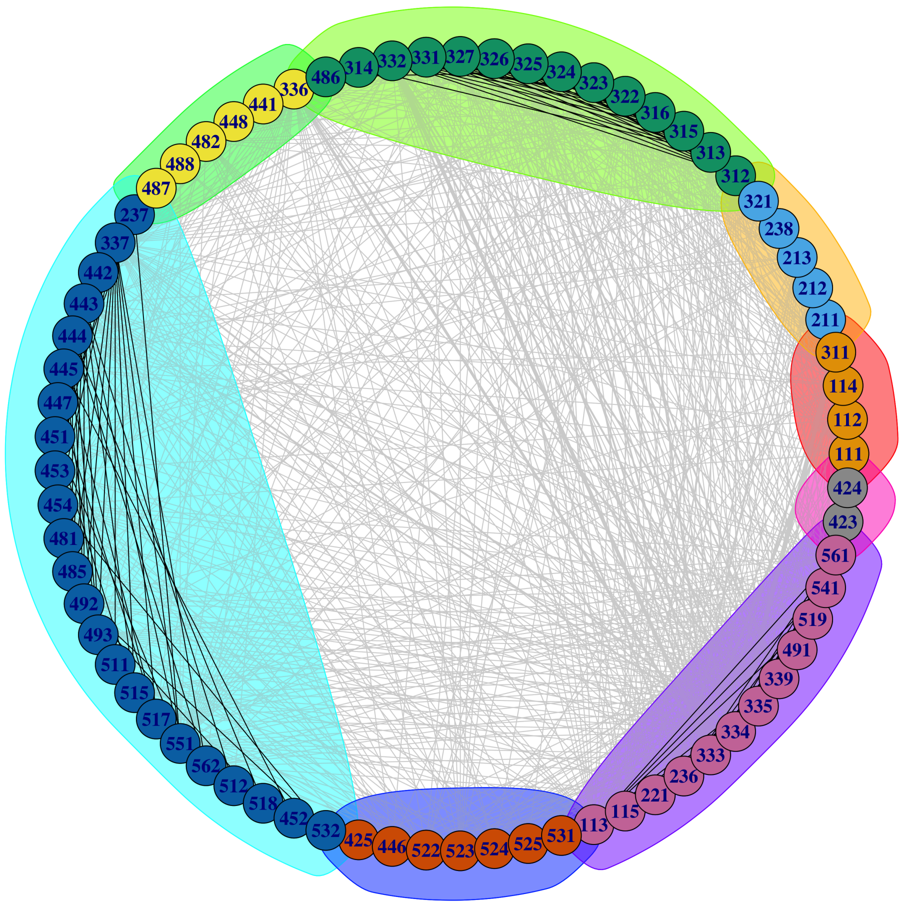
Community Detection and Resulting Knowledge Markets I use the rescaled undirected flow measure as a network edge weight to identify communities through the Louvain algorithm developed by Blondel et al. (2008). This procedure maximizes the modularity of the network by choosing the number of communities (knowledge markets) and the assignment of nodes (NAICS sectors) to communities. Modularity, a commonly used measure of connectedness of networks, measures the distance between the density of links *within* communities versus *between*.

This procedures produces 10 sets of NAICS 4-digit sectors that share the same inventors and have concentration measures. Applying the community detection algorithm results in knowledge markets that do not overlap: Each NAICS 4-digit sector belongs to one and only one knowledge market. Figure 2 displays the result of my procedure applied to NAICS 3-digit sectors. I report this exercise since the 4-digit equivalent would be too dense to depict. However, the knowledge markets identified by the two exercises are qualitatively similar although they are clearly more numerous in the 4-digit case. In this figure, lines denote inventor transitions, with width proportional to the effective undirected inventor flow between sectors. Nodes represent NAICS 3-digit. Black lines depict flows within knowledge markets, while gray lines represent transitions between communities.

Three features are worth emphasizing. First, the network is very dense, and transitions across 3-digit as well as 2-digit sectors are pervasive, differing largely in intensity. This approach is far more illuminating than grouping sectors based on broad product markets, which would neglect the linkages across disparate markets, or pooling all sectors together, which would neglect the difference in the strength of inventor flows. Second, the flows between communities appear more numerous than within communities, but this is solely a by-product of the circular layout of the network, whereby nodes mask flows within close communities on the circle. When applying the algorithm to 4-digit sectors, I find that less than a third of flows occur between communities, as expected since the community detection algorithm maximizes the density of within-community linkages. Third, and perhaps most importantly, the classification that I obtain sensibly groups together sectors that we might expect to share similar knowledge to innovate. Starting from sector 111 and going counter-clockwise, the knowledge markets in the figure can be described as follows. The first market, including sector 111, groups sectors involving agricultural production (111, 112 and 114) and food manufacturing (311). The second market, starting with 211, includes oil, gas, and mining. The green cluster at the top of the figure groups several heavy manufacturing industries, such as chemicals plastics and petroleum products, and pipeline transportation (486). The market in yellow consists largely of transportation services and manufacturing as well as motor vehicle dealers. The large blue cluster captures many retail sectors, as well as data processing, telecom, and broadcasting services. The remaining three markets include insurance and finance (red cluster), computer, electronics, and machinery manufacturing and professional services (violet), and wholesalers (gray).

Knowledge markets are identified using my measure of effective inventors, but the algorithm produces nearly identical results when using raw inventor counts; more than 97% of 4-digit NAICS sectors are classified in the same manner using the two measures. That is, 97% of sector pairs belong to the same knowledge market according to both measures.

Figure 2: Knowledge Markets Obtained from NAICS 3-digit Sectors



Note: This figure displays the network of inventor flows between NAICS 3-digit sectors and the knowledge markets resulting from the application of the Louvain community detection algorithm. Lines denote inventor transitions, with width proportional to the effective undirected inventor flow between sectors. Nodes represent NAICS 3-digit sectors. Black lines depict flows within knowledge markets, while gray lines transitions between communities.

2.3 Other Constructed Measures and Aggregation at Census Frequency

Patent Citation Measures For each patent classified by Goldschlag et al. (2016), I compute self-citations, forward citations, and a measure of patent generality. To count self-citations, I first identify the set of cited patents that belong to the same assignee as the citing patent. I weigh self-citations to account for cited patents that have multiple assignees. I count as one self-citation instances where the patent has a single assignee, and as one half if the cited patent has multiple assignees. The share of self-citations is given by the sum of weighted self-citations divided by the number of patents cited by each assignee. I construct five measures to correct self-citations for the assignee's importance in the relevant technology class of cited patents. For each citation made, excess self-citations are defined as

$1 - Pr$ (self-citation). The measure depends on how the probability of self-citation is computed. For the first three measures, I compute this probability as the assignee's share of total patents in the NAICS code attributed to the citing patent. I employ in turn the share of NAICS patents for the year of citation, the previous five years, and the cumulative share from the beginning of the sample. The other two measures are based on the CPC classification at the group and subgroup levels (the lowest levels of detail in the classification). For this measure, the probability of self-citation is derived for each citation by taking the share of patents by the assignee in the CPC (sub)group and the year corresponding to the cited patents.[6] Finally, I aggregate all measures across assignees in the same NAICS 4-digit code using the number of patents in the relevant code by each assignee in each year.

I also construct two truncation-corrected forward citation measures and a patent generality measure following the definitions and the procedures described in Hall et al. (2001). The forward citation measures compute the average number of citations received by each firm's patents, giving an indication of the importance of each patent for future technological developments. The correction for truncation is conducted by estimating the empirical CDF of the forward citations lag distribution of patents in the relevant CPC 2-digit technology class. The correction is then carried out by dividing the overall number of forward citations at the latest available date by the inverse of the CDF thus obtained. The procedure suggested by Hall et al. (2001) uses only information pertaining to the CPC 2-digit technology class of the cited patent. I also conduct an alternative correction that estimates a separate distribution for each citing CPC 2-digit class and sums the corrected forward citations across all citing classes. Patent generality also measures the technological impact of patents, but rather than focusing on citations it examines the scope of application of the patent. In particular, it measures the dispersion of citations received across different CPC classes. The higher the dispersion, the wider the technological applicability of the patent.[7]

Regulation Data Mercatus RegData provides a count of restrictions imposed on a number of NAICS 4-digit product markets, obtained by matching a set of keywords in NAICS descriptions to regulatory texts, and then taking the best match for each document. However, the available data does not include a set of codes due to data quality reasons.

Therefore, I process the description of NAICS 4-digit codes and compute the cosine-similarity between all pairs of sectors. I build an estimate of sector-relevant restrictions for missing sectors by taking an average weighted by cosine similarity of sectors included in RegData. I include in the average the five most similar NAICS codes if similarity is larger than .2, and I attribute the regulations of the most similar sector otherwise. I chose this threshold by inspecting the similarity associated to various NAICS pairs, and the assignment of regulations to sectors is not highly sensitive to this choice.

[6] This procedure is similar to the approach followed in Akcigit and Kerr's (2018) Appendix C.

[7] The interested reader should consult Acemoglu et al. (Forthcoming) for a detailed discussion, and the related appendix for details on the construction of these measures.

Inventor Distribution Measures I employ the measure of effective inventors constructed as detailed above to compute measures of researchers' concentration within sectors for each year in my sample. Specifically, I use the PatentsView assignee ID to identify firms that employ specific inventors in each sector, and then compute several measures of the concentration of inventors within sectors. I focus on the top 10% and bottom 50% share of inventors. I also use other common measures of dispersion like the ratio of the 90^{th} quantile to the median. I compute the Gini coefficient of inventors across CPC classes and NAICS 4-digit, assigning effective inventors to the relevant technology class or NAICS sector, to document increasing concentration of inventors in specific patent classes and sectors.

Aggregation at Census Frequency Data from the Economic Census are available at five-year intervals for the years 1997-2017, which requires aggregating the other data at the same frequency. Since I am interested in the effect of concentration on the allocation of inventors, I average all variables related to inventors and productivity using the five-year window *starting* in the census year (e.g., 1997- 2001 for 1997), while I use concentration measures for the corresponding census year. In the IV regression I use product restrictions as an instrument for concentration, which is why I average restrictions in the five-year window *ending* in the census year (e.g., 1993-1997 for 1997). Since Goldschlag et al. (2016)'s matching only covers the period up to 2016, I run all specifications in long-differences over the time frame 1997-2012.

3 Empirical Analysis

This section presents four main findings that apply to the period 1997-2012: (i) effective inventors became more concentrated across economic sectors; (ii) sectors with increased product market concentration attracted a growing share of relevant inventor types; (iii) growth in the share of relevant inventors negatively correlated with inventor productivity, as measured by forward citations as well as average growth in output per worker divided by effective inventors employed; and (iv) growth in the share of relevant inventors positively correlated with the share of self-citations and excess self-citations, as well as concentration of inventors at the top within sectors.

Results (i) and (ii) indicate a positive causal link between the growth in product market concentration and the increase in sectors' inventor share. Findings (iii) and (iv) point to misallocation: Inventor concentration in less competitive sectors turns out to be inefficient, as researchers are predominantly employed on projects that do not contribute to the growth of the sector. This work amounts to defensive innovation, as evidenced by the decline in forward citations of patents obtained by these firms and the decrease in growth per inventor that accompanies the increase in product market concentration.

The rest of this section proceeds as follows. The first subsection presents my empirical framework and variable definition. Remaining sections present in order results (i)-(iv) above. I discuss the causal interpretation of my results through an IV specification in Subsection 3.2.

3.1 Variable Definitions and Main Specification

Key to my analysis are measures of inventor concentration and of R&D productivity. I rely on the definition of effective inventors Î±i , that is productivity-adjusted inventors as explained in Section 2.2. For each product market p, I define the share of inventors employed by the sector in year t as

$$\text{Inventor Share}_{p,t} \equiv \frac{\sum_{p(i,t)=p} \alpha_i}{\sum_{k(i,t)=k} \alpha_i},$$

where the numerator represents the sum of effective inventors cited in patents registered in product market p, while the denominator consists of the total effective inventors that belong to the knowledge market. Effective inventors α_i are also the measure I use to evaluate the dispersion of inventors across sectors and technology classes. My results are robust to computing the inventor share using raw counts of researchers instead of effective inventors.

When analyzing R&D productivity I focus on the three patent-based measures described in Section 2.3, that is, forward citations, share of self-citations, and patent generality. Further, I compute a more direct measure of the productivity of inventors given by calculating the growth in output per worker divided by the number of effective inventors employed by the sector.

In most specifications, the independent variables are measures of concentration and controls for the size of the sector considered. As discussed in Section 2.1, my baseline measure of concentration is the lower bound of the Herfindal-Hirschman Index constructed by Keil (2017) using top sales share reported by the Economic Census for each sector.[8] I label this variable $\underline{\text{HHI}}_{p,t}$, where the line below stands for the lower bound. I chose this measure because my sample includes a relatively small number (about 80) of sectors that have an HHI index reported by the Economic Census. Using the lower bound allows me to expand the sample to157 sectors. The Economic Census HHI and its lower bound estimate are highly correlated and produce equivalent results, as shown in Table 2.

I obtain measures of sales from the Economic Census, which I deflate using BLS NAICS-specific price indexes. I use sales variables for two purposes. First, real sales in 2012 are the weight in my regressions. Second, I use the logarithm real sales as well as a quartic in real sales to control for changes in the size of sectors during my sample period. For the selected subset of sectors that reports the

[8] The expression used to obtain this measure is:

$$\underline{\text{HHI}}_{p,t} = 4\left[\frac{\text{CR4}_{p,t}}{4}\right]^2 + 4\left[\frac{\text{CR8}_{p,t} - \text{CR4}_{p,t}}{4}\right]^2 + 12\left[\frac{\text{CR20}_{p,t} - \text{CR8}_{p,t}}{12}\right]^2 + 30\left[\frac{\text{CR50}_{p,t} - \text{CR20}_{p,t}}{30}\right]^2,$$

where "CR{X}" denotes the concentration ratio, that is the share of sales, of the top X firms. This measure is a lower bound, and coincides with the actual HHI if the sector has less than 50 firms, and sales share are distributed equally in each of the top 0-4, 5-8, 9-20, 21-50 brackets. Keil (2017) reports a correlation of $\underline{\text{HHI}}$ with the actual index of 0.93.

number of companies, I also explore the robustness of my findings to controlling for sales per company, which provide a proxy for the average firm size in these sectors.

Given these definitions, my main specification is a sector-level long-difference regression over the period 1997-2012

$$\Delta \text{Share}_{p,\, 2012-1997} = f_k \mathbf{1}\left\{p \in k\right\} + \beta \Delta \underline{\text{HHI}}_{p,\, 2012-1997} + \gamma' \Delta \text{Size}_{p,\, 2012-1997} + \varepsilon_p, \tag{2}$$

where ΔShare denotes the change in the inventors' share of product market p; $f_k \mathbf{1}\left\{p \in k\right\}$ is a dummy variable that takes value 1 if the product market belongs to knowledge market k; $\Delta\underline{\text{HHI}}$ is the change in the HHI lower bound; and ΔSize is a set of controls for the size of sector p. Depending on the specification, ΔSize is the change in log real sales, the change in log real sales per firm, or the change in the terms of a quadratic polynomial in real sales.

Regressions are weighted by sector sales in 2012 for the findings which rely on Economic Census sector-level measures, and I estimate robust standard errors in all specifications. When looking at patent measures, I employ the same specification as in equation 2, where I replace the outcome variable with the change in patent productivity and the independent variable with the change in inventors' share. In this case, since I do not rely on Economic Census measures, I report unweighted regressions.[9] I also discuss the robustness of these findings to adopting the same specification using the HHI lower bound and weighting by sales.

3.2 Results

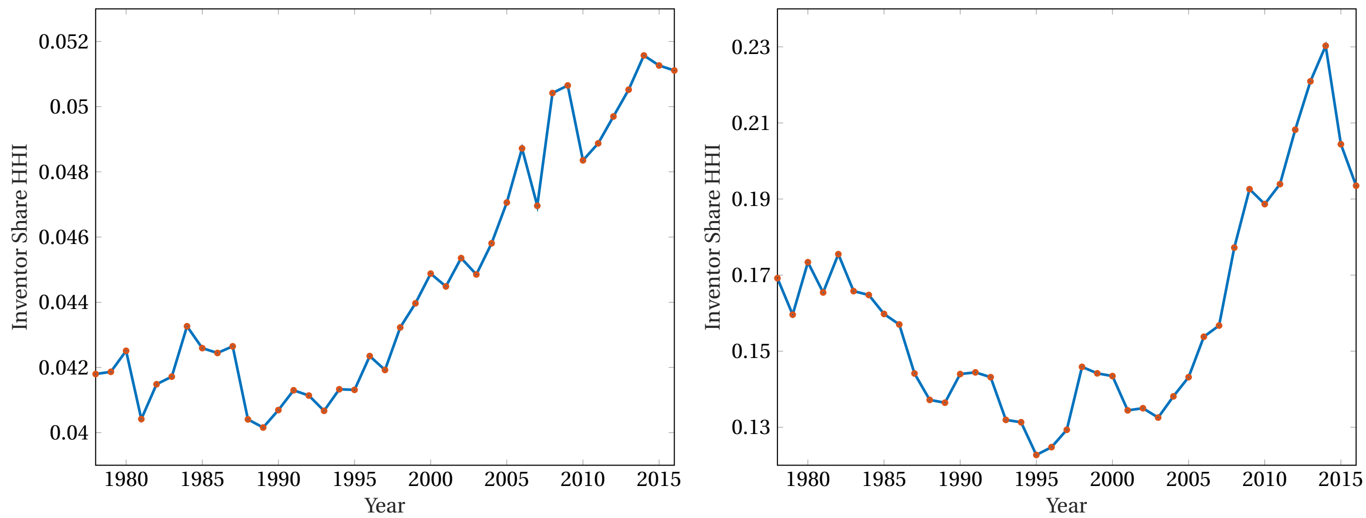
3.2.1 Inventor Concentration across NAICS Sectors has Increased

Figure 3 reports the time series of inventor concentration across NAICS 4-digit industries for the period 1976-2016, for which the Goldschlag et al. (2016) data is available. Panel (a) depicts the share of effective inventors and panel (b) that of raw inventors. I use the HHI index of inventor shares accruing to each sector as a measure of concentration. Both panels display an increasing concentration of inventors beginning in the late 1990s. These patterns align closely with trends reported in Akcigit and Ates (2019), which document a rising share of patents registered by top firms within sectors. Figure 3 extends those findings to the cross-industry allocation of inventors. The increase in inventor concentration is sizable, corresponding to about a 20% increase in the HHI for the effective inventor measure over the period 1997-2012. Based on raw inventor data, the figure rises to 30%.

As for the results presented below, the effective inventor measure and the raw inventor count behave similarly, although the series for raw counts is more volatile and exhibits larger changes. The

[9] As I will show, the change in the inventor share is highly correlated with the change in the HHI, so this specification essentially amounts to a rescaling of the coefficient that would be obtained using the HHI.

Figure 3: Herfindal-Hirschman Index of Effective Inventors across NAICS 4-digit Industries, 1976-2016



(a) Effective Inventors

(b) Inventor Count

Note: This figure reports the time series of inventor concentration, as measured by the HHI index of inventor shares across NAICS 4-digit sectors. The left panel reports the series constructed using effective inventors as defined in Section 2.2, the right panel uses instead raw inventor counts. Only the NAICS 4-digit sectors with data for all years are included.

figure for effective inventors is less volatile since this measure derives from a regression that residualizes time, firm, and technology class fixed effects.

3.2.2 Markets with Growing Concentration Increased Their Inventor Share

In this section, I present three sets of results for each specification, which differ in the estimation sample to account for extreme observations. In regression tables, "Full Sample" refers to the sample of observations with non-missing observations for all the variables included. I propose two sample selections to rule out that outliers drive the baseline results. "Trim Outliers" refer to a sample that excludes the most extreme observations for the outcome and the independent variable. I exclude the observations that fall beyond three standard deviations from the sample average of each variable and that are most likely to drive the results estimated using the full sample.[10] "Mahalanobis 5%" denotes the sample where I exclude the 5% extreme observations based on the Mahalanobis distance of pairs of observations from the data centroid. Since this procedure is based on the joint distribution of the outcomes and independent variables, the sample thus obtained varies in each regression.

Table [11] presents the results of regression (2) where the outcome variable is the change in knowledge-market inventor share, and the independent variable is either the change in the lower bound of the

[10] I justify the choices for each variable in detail in my replication code using the empirical kernel density and detailed tabulations.

[11] I justify the choices for each variable in detail in my replication code using the empirical kernel density and detailed tabulations.

Herfindal- Hirschman Index discussed above or that in the index as reported by the Economic Census. The results in Table 2 highlight a strong positive correlation between the change in the HHI and the change in the share of effective inventors accruing to each NAICS sector. Note that this regression is only partially driven by the contemporaneous correlation between the two variables. As discussed above, the share of effective inventors is averaged over the five years *starting* in the Economic Census year, while the concentration measures refer to the Economic Census year only.

Two important notes on the scale of the variables are in order. First, here and in all following tables and graphs, all variables that refer to shares or growth rates are reported in percentage points for ease of interpretation. With regard to the coefficient in Column (1) of Table 2, for example, an increase in one unit of the HHI index leads to an increase in the share of the relevant knowledge market of 27.25 percentage points. Second, HHI indices are constructed to range between 0 and 1. In 2012, the HHI lower bound has a sales-weighted average of .03 and a standard deviation of .032. According to Table 2, a standard deviation increase in this measure is associated with a .87 percentage point increase in the share of inventors accruing to the relevant NAICS sector. In comparison, the sales-weighted average share of inventors in 2012 is 1.18%, with a standard deviation of 1.82%, so the estimated effect of a one standard deviation increase in concentration corresponds to about half a standard deviation increase in the share of inventors in the relevant market. The estimates using the HHI lower bound tend to be noisier as this is a constructed, and therefore imprecise, measure of concentration. However, the number of available observations is much larger than the actual HHI, allowing me to extend my findings to about double the number of sectors.[12]

While suggestive, the correlation presented above neglects two fundamental components. First, it does not include controls for the size of the sectors or firms, which could have a confounding and mechanical effect on the share of scientists. Second, it estimates the correlation both across and within knowledge markets. In Table 3, I address these two limitations by restricting the analysis to within knowledge markets, and controlling for two measures of size. In the upper panel of Table 3, the change in the logarithm of real sales serves as a measure of the change in the size of each sector, while the lower panel shows the results when average sales per firm are included as a control. I include sales per firm to account for the fact that there might be significant barriers to entry in R&D. These barriers might be easier to overcome for larger firms, mechanically linking concentration and inventor hiring. Since the Economic Census reports the number of companies only for a subset of firms, the sample used in the lower panel is smaller than in the upper panel. The results in Table 3 confirm the positive relation between the change in inventor shares and concentration. They are largely unchanged relative to the estimates in Table 2, suggesting that the correlation does not arise mechanically from factors related to firm or sector size. In particular, these findings imply that sectors with increasing concentration have attracted a rising share of scientists above what would be predicted by their expansion in overall sales

[12] Regressions using the Economic Census HHI not reported in the main text or the Appendix are available on request.

and in average firm size.

Figure 4 depicts graphically the residualized observations underlying the estimated coefficients in Columns (2) and (6) of Table 3, Panel (a). The upper panel portrays changes of knowledge-market inventor shares over the change in the HHI lower bound, after partialling out fixed effects for the relevant knowledge market and changes in log real sales. The marker size is proportional to the regression weight. Although the sample displays some observations that appear extreme, the bulk of observations—and especially of weighted observations—falls on the regression lines, mitigating concerns that a few outliers might drive the results. In any event, I explore the robustness of the results to the exclusion of non-residualized observations, identifying extreme observations either manually, or using the Mahalanobis distance. Importantly, this exercise reveals that the observations that appear extreme in the residualized scatter are not unusual when considering the marginal or joint distribution of non-residualized outcome and independent variables. The bottom panel of Figure 4 reports the binned scatter plot corresponding to the sample where the 5% extreme observations according to the Mahalanobis distance have been removed. It confirms that the positive relation between concentration and inventor shares is not driven by a few extreme observations. The corresponding regression results in Table 3(a), Column (6), show that the estimated coefficient is significant at a 5% confidence level. The results presented in this section are robust to using the raw number of inventors to compute the share of researchers captured by each product market.

Appendix Table 11 shows estimates using the share of effective inventors of each product market over the total. This amounts to neglecting the fact that inventors flow only across sectors that can employ their skills. In this specification, I find a significant, albeit small, effect of product market concentration on the share of inventors. However, this result only arises when the sample is trimmed to remove outliers. This is not surprising, considering that mismeasuring the labor market for inventors should bias the estimates of inventor mobility towards zero, since many of the sectors would not be routinely connected by inventor flows. Addionally, this result conforms with the findings in Table 3, which show that including knowledge-market fixed effects does not alter the coefficients significantly, suggesting that flows across knowledge markets are indeed negligible.

Appendix **??** establishes the robustness of all the findings in this section to the use of raw inventor counts rather than effective inventors to compute both inventor shares and knowledge markets.

IV Results I now present instrumental variable results that suggest that the relation between concentration and inventor shares is causal. Indeed, more concentration could be the result of increasing technological entry barriers as incumbents hire more R&D inventors. In this scenario, the causality would flow from increased inventor shares to higher concentration. Above, I tried to mitigate this concern using as my outcome variable the average share of inventors following the Economic Census years to which the HHI refers. However, reverse causality could still be present if the autocorrelation of

Table 2: Regressions of Change in 4-digit Knowledge Market Share over Change in HHI Measures, Long-Differences, 1997-2012

	Δ Inventor Share (pp)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\underline{HHI}$	27.293*		27.183*		27.326*	
	(11.569)		(11.941)		(11.620)	
Δ HHI		22.399***		22.399***		22.350***
		(6.345)		(6.345)		(6.343)
Knowledge Market FE						
Sample	Full Sample	Full Sample	Trim Outliers	Trim Outliers	Mahalanobis 5%	Mahalanobis 5%
Weight	Sales	Sales	Sales	Sales	Sales	Sales
Observations	157	80	155	80	150	71

Note: Regressions weighted by sales in 2012; robust standard errors in parentheses; symbols denote significance levels $\left(+\ p < 0.1,^{*}\ p < 0.05,^{**}\ p < .01,^{***}\ p < .001\right)$; checkmarks indicate the inclusion of fixed effects. This table presents the results of specifications (2), when the outcome is the share of effective inventors of sector p over total inventors in knowledge market k, and the independent variable is the change in the lower bound of the Herfindal-Hirschman Index for product market p, as implied by Economic Census concentration ratios, or the HHI index reported in the Economic Census. "Full Sample", "Trim Outliers" and "Mahalanobis 5%" refer to the samples described in the main text.

Table 3: Regressions of Change in 4-digit Knowledge Market Share over Change in HHI Lower Bound, Long-Differences, 1997-2012

(a) Controlling for Change in Log Real Sales

	Δ Inventor Share (pp)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\underline{HHI}$	26.093*	22.509*	25.904*	22.716*	26.111*	22.554*
	(10.696)	(10.848)	(11.124)	(10.948)	(10.725)	(11.019)
Δ log Sales	0.914**	0.548*	0.881**	0.539*	0.918**	0.562*
	(0.278)	(0.243)	(0.275)	(0.242)	(0.283)	(0.261)
Knowledge Market FE		✓		✓		✓
Sample	Full Sample	Full Sample	Trim Outliers	Trim Outliers	Mahalanobis 5%	Mahalanobis 5%
Weight	Sales	Sales	Sales	Sales	Sales	Sales
Observations	157	153	155	152	150	139

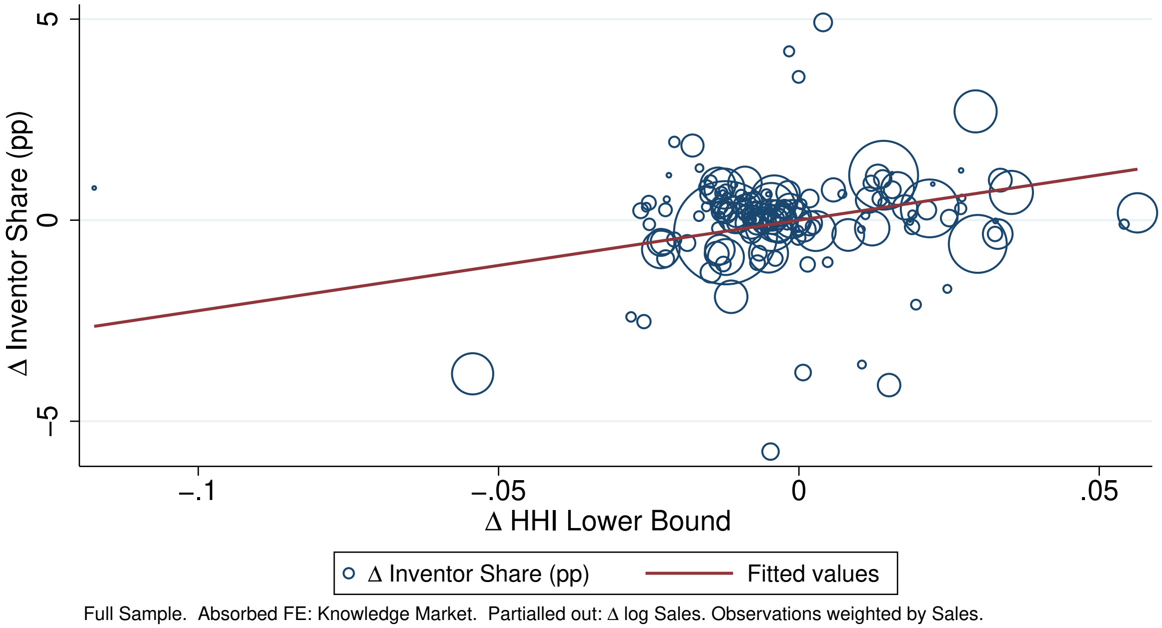
(b) Controlling for Change in Log Real Sales per Company

	Δ Inventor Share (pp)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\underline{HHI}$	35.230**	20.783+	35.230**	20.783+	35.154**	22.854*
	(12.759)	(10.615)	(12.759)	(10.615)	(12.647)	(11.197)
Δ log Size	0.175	-0.040	0.175	-0.040	0.300	-0.055
	(0.382)	(0.253)	(0.382)	(0.253)	(0.460)	(0.346)
Knowledge Market FE		✓		✓		✓
Sample	Full Sample	Full Sample	Trim Outliers	Trim Outliers	Mahalanobis 5%	Mahalanobis 5%
Weight	Sales	Sales	Sales	Sales	Sales	Sales
Observations	81	79	81	79	75	67

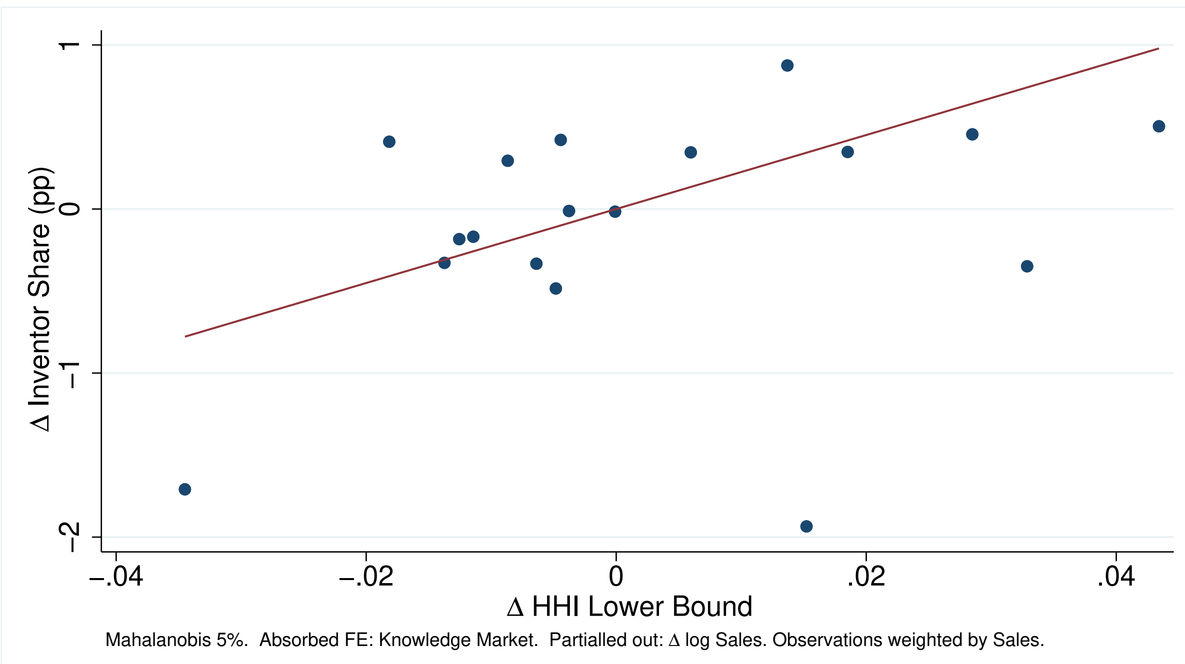
Note: Regressions weighted by sales in 2012; robust standard errors in parentheses; symbols denote significance levels $\left(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001\right)$; checkmarks indicate the inclusion of fixed effects. This table presents the results of specifications (2), when the outcome is the share of effective inventors of sector p over total inventors in knowledge market k, and the independent variable is the change in the lower bound of the Herfindal-Hirschman Index for product market p, as implied by Census concentration ratios. "Full Sample", "Trim Outliers" and "Mahalanobis 5%" refer to the samples described in the main text.

Figure 4: Residualized Scatter Plots Corresponding to Selected Columns in Table 3, Panel (a)

(a) Raw Scatter Plot, Specification in Column (2)



(b) Binned Scatter Plot, Specification in Column (6)



Note: This figure presents residualized scatter plots of the change in the share of effective inventors of sector p over total inventors in knowledge market k, over the change in the lower bound of the Herfindal-Hirschman Index for product market p, as implied by Census concentration ratios. The upper panel reports the data for the full sample, where both variables are residualized by change in log real sales and knowledge market fixed effects. The size of the markers is proportional to the weight of each observation in the regression (sector sales in 2012). The regression line uses the coefficient on the change in HHI lower bound in Column (2) of Table 3. The lower panel presents a binned scatter plot removing the observations with the highest 5% Mahalanobis distance from the sample centroid. Observations are aggregated using sales weights and the regression line is from Column (6) of Table 3.

inventor shares is sufficiently high. As a consequence, I have calculated 2SLS estimates that instrument the change in the HHI lower bound with changes in product market restrictions, as measured by the Mercatus dataset RegData 4.0. Theoretically, an increase in restrictions should raise barriers to entry in affected product markets, thus leading to higher concentration. As discussed below, such proves to be the case empirically, validating sector-specific restrictions as an instrument for concentration. A violation of the exclusion restriction requires a causal connection between product market regulations and the share of inventors hired by each sector, independent of product market concentration. For example, regulations affecting existing technologies might require more inventors to meet product market restrictions. However, this effect is unlikely to be sufficiently large and persistent to be captured by my measure of inventor shares. Further, the regulations counted in RegData are not exclusively product restrictions, but also include reporting obligations and other legal burdens that are not related to technological components. In addition, while product restrictions might certainly induce a change in the direction of innovation, there is no a priori reason to believe that the scale of innovation activity should increase. These considerations lead me to believe that the exclusion restriction is not likely to be violated.

The results of the 2SLS estimation are presented in the upper panel of Table 4. The specification is the same as in Column (2) of 3, including both knowledge market and sale change fixed effects. The 2SLS estimates confirm the significance of concentration changes for the increase in knowledge market inventor shares. The magnitudes of estimated coefficients are statistically indistinguishable from the ones reported in the baseline regression. The first-stage F clearly indicates that instruments are weak. This is unsurprising since, as detailed above, both the HHI lower bound and the regulation measures are estimated. In particular, I had to impute regulations for a large part of the sample using the cosine-similarity between product market restrictions.[13] However, instruments are not irrelevant. The results in the lower panel of Table 4 imply that the first-stage t-statistic for the regression of the change in the HHI lower bound over log-regulations is 2.07, which corresponds to a p-value of 0.041. The reduced form regression of inventor share over log restriction change is as highly significant. Accordingly, the SW underidentification test rejects the null hypothesis at a 5% confidence level. Given the weakness of the instruments, I also report the Anderson-Rubin p-value and the corresponding confidence intervals in brackets, which confirm that the coefficient is 5% significant.

Taken together, the results presented in this section establish a causal link between the increase in inventor concentration and the shifts in product market concentration across NAICS 4-digit sectors.

[13] Using only available sectors requires dropping two thirds of the observations. See Appendix A for details on data construction.

Table 4: IV Regressions of Change in 4-digit Knowledge Market Share over Change in HHI Lower Bound, 2SLS Long-Difference, 1997-2012

(a) 2SLS Results

	Δ Inventor Share (pp)	
	(1)	(2)
$\Delta \underline{HHI}$	32.426+	30.096+
	(16.987)	(15.819)
	[4.850, 99.013]	[4.415, 92.104]
Δ log Sales		0.525*
		(0.247)
		[0.525, 0.525]
Knowledge Market FE	✓	✓
Sample	Full Sample	Mahalanobis 5%
Weight	Sales	Sales
Observations	157	150
First-Stage F	4.65	4.75
Anderson-Rubin p-value	.0298	.0321

(b) First Stage and Reduced Form

	Δ Inventor Share (pp)	$\Delta \underline{HHI}$
	(1)	(2)
Δ log Restrictions	0.478*	0.016*
	(0.220)	(0.007)
Δ log Sales	0.539+	-0.000
	(0.274)	(0.005)
Knowledge Market FE	✓	✓
Sample	Full Sample	Full Sample
Weight	Sales	Sales
Observations	153	153

Note: Regressions weighted by sales in 2012; robust standard errors in parentheses; symbols denote significance levels $\left(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001\right)$; checkmarks indicate the inclusion of fixed effects. This table presents the results of specifications (2), when the outcome is the share of effective inventors of sector p over total inventors in knowledge market k, and the independent variable is the change in the lower bound of the Herfindal-Hirschman Index for product market p, as implied by Economic Census concentration ratios, instrumented by the change in log-restrictions relevant to the NAICS sector. The lower panel present first-stage and reduced-form relations. "Full Sample" and "Mahalanobis 5%" refer to the samples described in the main text.

3.2.3 Sectors that Attracted More Researchers Saw Increasing Top Firms' Inventor Shares and Falling Patent Forward Citations

While the findings presented so far establish a connection between inventor and product market concentration, they do not establish that changes in the distribution of researchers across sectors are inefficient. It would not be unreasonable, for example, to think that more concentrated sectors saw increased entry as a result of the higher rents captured by incumbents. Table 5 shows that the opposite occurred. Specifically, the share of effective inventors accruing to top inventor-hiring firms increased in the sectors that attracted more inventors over the period considered, relative to firms with fewer inventors in the sector—a finding consistent across a variety of measures displayed in Columns (1) to (6). These outcomes suggest that inventors have increasingly concentrated among large incumbents, that is, sectors that increased their inventor share also saw a *within-sector* increase in inventor concentration.

Throughout this section, I present results using changes in inventor shares to focus directly on the correlation between inventor transitions and their within-sector distribution. Unless otherwise noted, these findings are robust to using the change in the HHI rather than the inventor share, as should be expected from the strong correlation between these two variables reported in previous tables. For this section, and other patent-based measures, I present robustness results using the change in the HHI in Appendix B.4.

My next finding suggests that inventor concentration is driven by a rise in defensive innovation, that is R&D aimed at protecting the incumbents' dominant position and raising barriers to entry. Table 6 shows that inventors' concentration in specific sectors went hand in hand with a fall in forward citations for patents, a standard measure of a patent's contribution to further innovations (Hall et al., 2001). The result in Columns (1) and (2) report two different measures of forward citations that differ in how the series are corrected for truncation. As discussed in Section 2.3, the measure in Column (2) uses the procedure delineated by Hall et al. (2001), computing the forward citation lag distribution conditioning on the technology class of the cited patent. Column (2) also conditions on the technology class of citing patents. Column 3 presents the estimates relative to patent generality, a measure of patent impact that increases with the scope of application. The regressions in this table are unweighted since they rely only on patent data, but results are robust to using the HHI as a regressor and weighting by sales. I present results for the full sample, as well as restricting to the middle range of changes in inventor shares, which contains more than 90% of the observations. In both samples, I find a highly significant negative relation between changes in inventor shares and the fall in forward citations. The coefficients imply a high semi-elasticity of self citations to changes in the inventor shares, whereby a 1 percentage point increase in the share of inventors leads to a 0.2-0.5% reduction in forward citations. After dropping extreme observations, I also find a significant decrease in the generality of the patents, indicating that concentrating sectors produce less widely applicable patents. However, the generality finding is not robust to estimating the regression using the HHI as the independent variable.

The fall in forward citations is a first indication of the presence of defensive innovation (see, e.g., Guellec et al., 2012). In the next section, I show that these patents also appear to do relatively little to boost productivity, as measured by growth in output per worker.

Before moving to the results on productivity, I investigate a competing explanation for my findings on output growth. As highlighted by Acemoglu et al. (Forthcoming) and Akcigit and Kerr (2018) among others, large incumbents have a strong incentive to focus on improving their own products at the expense of broadly applicable innovation. This mechanism would also imply that an increase in incumbents' share of R&D resources leads to falling innovation productivity. In order to assess the importance of this channel, and in keeping with the analysis in Akcigit and Kerr (2018), I use the share of self-citations to measure the extent of internal innovation conducted by firms. Table 7 displays the results pertaining to this measure. All columns use as dependent variable the change in excess log self-citations as defined in Section 2.3. Columns (1) and (2) build excess self-citations correcting for the importance of firms' patents for the CPC group, which reflects the technological classification of the patent. Columns (3) and (4) use the more narrowly defined CPC subgroups for robustness. Coefficients are mostly non-significant and turn negative when knowledge market fixed effects are included. Column (3) displays a marginally significant coefficient. However, this result is not robust to using the HHI as regressor and weighting regressions by sales as in the baseline specification. The findings in this table suggest that incremental innovation does not drive my results.

Table 5: Regressions of Change in Inventor Distribution Measures over Change in 4-digit Knowledge Market Share, Long-Difference, 1997-2012

	Δ 90/50 Quantile Ratio	Δ Top 10%/Bottom 50%	Δ Top-50/Bottom-50 Share Ratio	Δ Top 10%	Δ Bottom 50%
	(1)	(2)	(3)	(4)	(5)
Δ Inventor Share (pp)	0.211+	0.243*	0.314+	0.018**	-0.008*
	(0.107)	(0.097)	(0.184)	(0.006)	(0.004)
Δ log Sales	-0.100	0.328	0.147	0.026	0.005
	(0.122)	(0.294)	(0.316)	(0.020)	(0.007)
Knowledge Market FE	✓	✓	✓	✓	✓
Sample	Full Sample	Full Sample	Full Sample	Full Sample	Full Sample
Weight	Sales	Sales	Sales	Sales	Sales
Observations	118	118	118	118	118

Note: Regressions weighted by sales in 2012; robust standard errors in parentheses; symbols denote significance levels $\left(+\ p < 0.1,^{*}\ p < 0.05,^{**}\ p < .01,^{***}\ p < .001\right)$; checkmarks indicate the inclusion of fixed effects. Please refer to notes in Table 3 for further details. Column (1) uses the ratio in the 90 percentile of effective inventors to the median as the outcome variable. Columns (2) and (3) instead present the share ratio, that is the share of effective inventors accruing to the top 10 or 50% relative to the share accruing to the bottom 50% of the distribution within each NAICS sector.

Table 6: Regressions of Changes in Forward Citation over 4-digit Knowledge Market Share, Long-Differences, 1997-2012

(a) Full sample

	Δ log Citations/Patent (CPC)	Δ log Citations/Patent (Total)	Δ Patent Generality
	(1)	(2)	(3)
Δ Inventor Share (pp)	-0.197***	-0.227***	-0.004
	(0.044)	(0.051)	(0.004)
Δ log Sales	-0.234*	-0.258+	0.008
	(0.112)	(0.148)	(0.013)
Knowledge Market FE	✓	✓	✓
Sample	Full Sample	Full Sample	Full Sample
Weight			
Observations	153	153	153

(b) Full sample, restricting to the middle range of the change in inventor shares (−2% to +2%)

	Δ log Citations/Patent (CPC)	Δ log Citations/Patent (Total)	Δ Patent Generality
	(1)	(2)	(3)
Δ Inventor Share (pp)	-0.545***	-0.618***	-0.025*
	(0.113)	(0.137)	(0.012)
Δ log Sales	-0.232*	-0.255+	0.008
	(0.109)	(0.146)	(0.012)
Knowledge Market FE	✓	✓	✓
Sample	Full Sample	Full Sample	Full Sample
Weight			
Observations	144	144	144

Note: Unweighted regressions; robust standard errors in parentheses; symbols denote significance levels $\left(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001\right)$; checkmarks indicate the inclusion of fixed effects. This table present the results of specification (2), when the outcome is the log-change in forward citations and the change in patent generality in sector p over the change in the share of inventors employed in sector p. Column (1) and (2) presents the results when forward citations are extrapolated the procedure Hall et al. (2000) to avoid truncation bias. A specific cite-lag distribution over 35 years is estimated for each pair of cited and citing CPC2-codes. Column (1) employs the extrapolation scheme by each pair of CPC2 cited and citing sector. Column (2) applies the extrapolation scheme to total citations received by each cited patent. Column (3) presents results on the patent generality measures. All columns exclude self-citations. Upper panel: full sample; bottom panel: excluding sectors with absolute increase in the inventor share above 2%.

Table 7: Regressions of Change in Excess Self-Citations over 4-digit Knowledge Market Share, Long-Differences, 1997-2012

	Δ CPC group self-citations		Δ CPC subgroup self-citations	
	(1)	(2)	(3)	(4)
Δ Inventor Share (pp)	0.920	-0.444	0.958+	-0.228
	(0.711)	(1.083)	(0.512)	(0.801)
Δ log Sales	-1.841	-1.954	-1.456	-1.674
	(1.925)	(1.988)	(1.326)	(1.279)
Knowledge Market FE		✓		✓
Sample	Full Sample	Full Sample	Full Sample	Full Sample
Weight				
Observations	157	153	157	153

Note: Unweighted regressions; robust standard errors in parentheses; symbols denote significance levels $(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001)$; checkmarks indicate the inclusion of fixed effects. This table presents the results of specifications (2), when the outcome is the change in excess self-citations in sector p over the change in the share of inventors employed in sector p.

3.2.4 Markets with Growing Inventor Shares Experienced a Fall in Inventor Productivity

Table 8 presents the results of running (2) when the outcome is the average growth in output per worker per effective inventor. I use growth in annual output per worker provided by the Economic Census and average this measure over the five-year window starting in the Economic Census year, and I analogously build a measure of average effective inventors over the same period. Inventor productivity is then defined as average output per worker growth divided by average number of effective inventors. Both the outcome and the dependent variable are measured in percentage points. Table 8 reveals a negative and significant correlation between the increase in the number of effective inventors and inventor productivity, regardless of the independent variable employed and the sample restriction adopted.

Starting from the upper panel of Table 8, the median change in the share of effective inventors over the period was .014pp, while the measure of effective inventors has a median of 2018.[14] The coefficient for concentration in Column (5) implies a fall of .15pp .15pp $(-.005 \times .014pp \times 2018)$ in average annual labor productivity growth. This number decreases to -. 28pp when considering only sectors with positive growth in labor productivity, which accounted for the bulk of the increase in inventor shares. An alternative back-of-the-envelope computation, using the change in product market concentration to predict the change in inventor shares gives even starker results. Using the coefficient in Column (2) of Table 3(a) and considering a median change in the HHI of 0.002 yields an increase in the share of effective inventors in concentrating sectors of 0.045pp. This implies a fall in average labor productivity implied by misallocation of 0.45pp. While these numbers might appear large considering the entirety of the economy, it is worth noting that my sample includes mainly manufacturing and retail sectors,

[14] Recall that effective inventors in each year are measured as the sum of inventor fixed-effects in each year, and therefore do not represent the simple count of inventors.

which experienced a sizable reduction of about 2.8pp in average annual productivity growth from 1997 to 2012, driven by a steep decline in output per worker growth in manufacturing. Therefore, the mechanism I propose would explain around 15 percent of the observed decrease in output per worker growth in these sectors.

The estimates in the lower panel of Table 8, which uses the HHI instead the change in inventor shares as independent variable, imply even larger growth effects. Using the estimates in Column (2), a median HHI change of 0.02 and a median number of effective inventors of 1421 in sectors with growth in inventor shares implies a −0.78pp change in output per worker growth from misallocation, with a confidence interval ranging from −0.13 to −1.45pp. The midpoint of these estimates would explain 27% of the observed fall in output per worker growth over the sample period, with bounds ranging from around 5% to about 50%.

This last set of results further supports the hypothesis that defensive innovation increased in concentrating sectors. To protect their dominant position, firms engage in such R&D to thwart innovation by potential competitors. Similarly citing defensive innovation as a motive, Argente et al. (2020) report that incumbent firms tend to register a large number of patents, but account for a small share of overall innovations.

Table 8: Regressions of Changes in Inventor Productivity over Changes in Inventors' Share and HHI, Long-Difference, 1997-2012

(a) Change in Inventors' Share as Independent Variable

	Δ Growth/Inventor (pp)			
	(1)	(2)	(3)	(4)
Δ Inventor Share (pp)	-0.007**	-0.005*	-0.007**	-0.005*
	(0.002)	(0.002)	(0.002)	(0.002)
Δ log Sales		-0.051*		-0.054*
		(0.021)		(0.021)
Knowledge Market FE	✓	✓	✓	✓
Sample	Full Sample	Full Sample	Mahalanobis 5%	Mahalanobis 5%
Weight	Sales	Sales	Sales	Sales
Observations	101	101	96	93

(b) Change in HHI as Independent Variable

	Δ Growth/Inventor (pp)			
	(1)	(2)	(3)	(4)
Δ$\underline{\text{HHI}}$	-0.332**	-0.292*	-0.332**	-0.290*
	(0.113)	(0.123)	(0.114)	(0.126)
Δ log Sales		-0.052*		-0.053*
		(0.021)		(0.022)
Knowledge Market FE	✓	✓	✓	✓
Sample	Full Sample	Full Sample	Mahalanobis 5%	Mahalanobis 5%
Weight	Sales	Sales	Sales	Sales
Observations	101	101	98	94

Note: Regressions weighted by sales in 2012; robust standard errors in parentheses; symbols denote significance levels $(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001)$; checkmarks indicate the inclusion of fixed effects. Please refer to notes in Table 3 for further details. Inventor productivity is measured as the average growth in output per worker over the five years starting in the Economic Census year over the total number of effective inventors in each sector. The upper panel presents estimates when the independent variable is the change in the share of inventors accruing to a sector, while the bottom panel uses the change in the lower bound of the HHI index.

4 Model

This section presents a Schumpeterian model based on Abrams et al. (2013), featuring growth through creative destruction by entrants, as well as the possibility for incumbent monopolist of researching a defensive technology that increases research costs for entrants. I first present a single-sector model to clarify the mechanism at play within each sector in the economy and study the properties of a constant-growth equilibrium analytically. I derive sufficient conditions under which markup increases lower R&D productivity, and show that this occur only if the distribution of inventors shifts in favor of incumbent firms carrying out defensive innovation.[15] I also show that in this model, when the

[15] In the empirical analysis, I used the HHI as a measure of concentration and market power. Appendix B.5 shows that the Lerner Index from NBER-CES, a standard measure of markups, is strongly correlated with the HHI in my sample, justifying the reduced-form mapping I adopt in this section.

supply of inventors is rigid, inventors' productivity is unaffected by markup changes. Then, I move to consider a two-sector model, where each sector is identical to the single-sector model, and the supply of inventors is perfectly rigid, which shuts down within-sector misallocation occurring independently of inventors' movements across sectors. I show that increasing markups in one of the two sectors of the economy lead to a misallocation of R&D resources towards defensive innovation in the less competitive sector. Finally, I study the optimal allocation of R&D subsidies needed to achieve maximum growth in a calibration of the two-sector model that matches moments of the R&D distribution in 1997, the starting year for my empirical analysis.

All omitted proofs are reported in Appendix C.

4.1 Single-sector Model

4.1.1 Preferences and production

Consider the following continuous time economy with a single final good. There is a representative household with King-Plosser-Rebelo preferences over consumption and R&D labor

$$\mathbb{E}_t \int_t^{\infty} \exp\left(-\rho\left(s-t\right)\right)\left(\ln C_s - \frac{\chi\left(L_s^{RD}\right)^{1+\frac{1}{\phi}}}{1+\frac{1}{\phi}}\right)\mathrm{d}s, \tag{3}$$

where ϕ is the Frisch labor supply. In addition, the representative household inelastically supplies L units of production labor.[16] The representative household owns a differentiated portfolio of all the firms in the economy, with rate of return r_t, and receives a wages w, w^{RD}, for each unit of production and research labor, respectively. I assume that the economy is closed and that the final good is only used for consumption, $C_t = Y_t$. The above utility function yields the Euler equation

$$\frac{\dot{C}_t}{C_t} = r_t - \rho.$$

It also yields the R&D labor supply

$$L_t^{RD} = \left(\frac{w_t^{RD}}{\chi C_t}\right)^{\phi}.$$

The consumption good in the economy in each instant is a Cobb-Douglas aggregate of a measure-one continuum of products

$$\ln Y_t = \int_0^1 \ln y_t(i)\mathrm{d}i. \tag{4}$$

The consumption good in the economy is taken as the numeraire. The market for each product $y_t(i)$

[16] While this assumption is not necessary for the results to hold, it simplifies the analysis considerably. In the following section, I will consider both production and research labor as given by a fixed endowment in the constant growth equilibrium of the economy. In that case, the assumption is equivalent to assuming that both labor endowments grow at a constant rate.

consists of an incumbent and a fringe of competitors. In what follows, I focus on a single market, dropping the argument i. Each agent j in the sector has the linear production technology:

$$y_{j,t} = c_{j,t} l_{j,t},$$

where $c_{j,t}$ denotes the labor requirement for agent j to produce a unit of output, and $l_{j,t}$ denotes the production labor employed by the firm. The incumbent and competitors produce undifferentiated goods, and differ in their labor requirement. Competitors have labor requirement $c_{e,t} = c_t$, while the incumbent faces a lower unit labor requirement $c_{I,t} = \frac{c_t}{\phi}$, with $\phi > 1$. The incumbent maximizes profits by choosing a price p_t for her product. Profit maximization gives an optimum limit price $p_t = w_t c_t$, which leads her to capture the entire market and realize profits

$$\Pi_t = \left(c_{e,t} - c_{I,t}\right) w_t y_t = \left(\frac{\phi - 1}{\phi}\right) c_t w_t y_t.$$

Therefore, the incumbent acts as a monopolist, charging a markup $\phi > 1$ on its marginal cost. By the Cobb-Douglas assumption on the final good, the demand facing each product line is

$$y_t = \frac{Y_t}{p_t} = \frac{Y_t}{w_t c_t}.$$

Therefore, equilibrium normalized profits read

$$\frac{\Pi_t}{Y_t} \equiv \pi = \left(\frac{\phi - 1}{\phi}\right).$$

4.1.2 Innovation

Both incumbents and entrants can conduct innovation activity that, if successful, reduces their unit costs to

$$c_{I,t+\Delta_t} = \frac{c_{e,t}}{(1+\eta)\phi}, \quad \eta > 1$$

Here, η parametrizes the percentage increase in productivity for the innovating firm, relative to the technology previously operated by the incumbent. I assume spillovers from realized innovations as follows. Whenever either the incumbent or an entrant realize an innovation, all other firms gain access to a technology with unit costs

$$c_{e,t+\Delta_t} = \frac{c_{e,t}}{(1+\eta)}.$$

These assumptions imply that, if entrants realize an innovation, they outcompete previous incumbents and become the new monopolists. Displaced incumbents join the pool of entrants and from instant $t+\Delta t$ onwards operate the technology $c_{e,t+\Delta_t}$. This amounts to assuming that the incumbent's technology becomes obsolete after displacement.[17] Technically, this innovation structure removes the need to keep track of the number of realized innovations. Indeed, in each product market the relative productivity of incumbents relative to entrants is fixed at ϕ, making it possible to formulate the choice of innovation as a recursive problem.

Incumbents' and entrants' innovation differ in two respects. First, successful incumbents' R&D produces a *patent wall* of size $\omega > 1$, which decreases the success probability of entrants' innovations. Second, successful entrants' R&D results in an implemented innovation with certainty, while incumbents adopt new technologies with probability $\lambda \in [0,1]$. This parameter can be interpreted in two ways. The first interpretation is that λ captures the probability that the newly discovered technology is compatible with the incumbents' current production techniques. The second interpretation is related to the radical nature of incumbents' innovations. A low value for λ reduces the expected productivity improvement from an innovation by the incumbent. In other words, the lower λ, the more incremental are incumbents' innovations. Under these assumptions, incumbents *always* obtain a patent wall, but they only implement their innovations with probability λ.[18]

Following Acemoglu and Akcigit (2012), I assume that innovation investments consist in the choice of an arrival rate of new discoveries x_I, and that R&D costs are increasing and convex in this arrival rate. I specify incumbents' R&D costs as

$$C\left(x_I; w^{RD}\right) = \alpha_I \frac{x_I^{\gamma}}{\gamma} w^{RD},\ \gamma > 1,$$

where the term $\alpha_I \frac{x_I^{\gamma}}{\gamma}$ indicates the amount of inventors that the incumbent needs to obtain innovations with a flow probability x_I, and w^{RD} is the wage paid to inventors. For simplicity, I assume that incumbents can only have one available innovation at a time. That is, incumbents can only erect *one* patent wall of size $\omega > 1$, and cannot invest in further innovation until this wall is destroyed at a rate δ, which captures the pace of patent expiration.

Under these assumptions, incumbents' values at any given instant are just a function of the state of the patent wall in the product market they operate, $\Omega \in \{1, \omega\}$. Given the recursive nature of the problem, I drop time indexes in what follows. Incumbents' value functions read:

$$rV(1) - \dot{V}(1) = \max_{x_I} \left\{ \left(\frac{\phi - 1}{\phi}\right) Y - \alpha_I \frac{x_I^{\gamma}}{\gamma} w^{RD} + x_I \left(V(\omega) - V(1)\right) - x_{e,1} V(1) \right\}, \tag{5}$$

$$rV(\omega) - \dot{V}(\omega) = \left(\frac{\phi - 1}{\phi}\right) Y + \delta \left(V(1) - V(\omega)\right) - x_{e,\omega} V(\omega), \tag{6}$$

[17] An alternative interpretation of this assumption is that incumbents are forced to scrap the assets needed to operate the innovative technology upon destruction, so that the technology is no longer available to them onwards.

[18] This specification connects to the empirical evidence in Argente et al. (2020), who show that large incumbent firms tend to produce many patents, but implement a small number of product innovations.

where $x_{e,1}$ and $x_{e,\omega}$ denote entrants' innovation intensities, r is the interest rate in the economy to be determined in equilibrium, and δ the rate of patent expiration. The first line displays the flow value to incumbents that operate in a market not protected by a patent wall. There, incumbents realize instantaneous profits $\left(\frac{\phi-1}{\phi}\right)Y$, and choose their innovation intensity x_I, taking the researchers' wage w^{RD} and the entrants' innovation intensity $x_{e,1}$ as given. If entrants are successful at rate $x_{e,1}$, incumbents are destroyed. If incumbents' innovation is successful at rate x_I, they obtain the patent wall ω, which grants them the protected value $V(\omega)$. When a patent wall is in place, incumbents realize the same flow of profits as in the unprotected state, since economy-wide spillovers imply that incumbents are unable to reap profits from implemented innovations. However, incumbents face a different entrants' innovation intensity, $x_{e,\omega}$, which is lower than $x_{e,1}$ due to the patent wall in place, as I will show below. Finally, incumbents in state ω face a flow probability δ that the patent wall is exogenously destroyed, in which case they transition back to the unprotected state. Under these assumptions, the optimal incumbent's R&D decision is given by

$$x_I = \mathbf{1}\left\{V(\omega) - V(1) > 0\right\}\left(\frac{V(\omega) - V(1)}{\alpha_I w^{RD}}\right)^{\frac{1}{\gamma-1}}. \tag{7}$$

Following Abrams et al. (2013), I assume that each market has a mass of atomistic entrants, indexed by j, who face innovation costs that feature congestion externalities

$$C\left(x_{e,\Omega,j}; w^{RD}\right) = \zeta\Omega x_{e,\Omega,j} x_{e,\Omega} w^{RD}.$$

In this specification, ζ parametrizes the inventor requirement to obtain a unit aggregate entrants' innovation rate when the market is not protected by patent walls, $\Omega = 1$. Individuals costs are linear in the total entrants' research intensity in the product market, $x_{e,\Omega} \equiv \int_{\mathcal{J}} x_{e,\Omega,j}\,\mathrm{d}j$. In other terms, individual entry costs increase with the aggregate entry rate. Successful entrants obtain a new unprotected technology, regardless of the state of the market that they target, based on the assumption that they do not rely on patents from previous incumbents.

Under the above assumptions, the free entry condition for entrants targeting a market with patent wall Ω reads

$$\max_{x_{e,\Omega,i}}\left\{x_{e,\Omega,i} V(1) - \zeta\Omega x_{e,\Omega,i} x_{e,\Omega} w^{RD}\right\}.$$

This condition pins down the entry rate for each product market with patent wall Ω as

$$x_{e,\Omega} = \frac{V(1)}{\zeta\Omega w^{RD}}, \quad \Omega \in \{1, \omega\}. \tag{8}$$

This expression clarifies the effect of defensive innovation in this model. The size of the patent wall ω

represents the factor decrease in the entry rate when a defensive innovation is successful.

4.1.3 Equilibrium with Constant Growth

The laws of motion of product markets across protected and unprotected states is given by:

$$\dot{\mu}_1 = -\left(x_I + x_{e,1}\right)\mu_1 + \delta\mu_\omega + x_{e,\omega}\mu_{e,\omega} + x_{e,1}\mu_{e,1}, \tag{9}$$

$$\dot{\mu}_\omega = -\left(x_{e,\omega} + \delta\right)\mu_\omega + x_I\mu_1, \tag{10}$$

where $\mu_{e,\omega}$ and $\mu_{e,1}$ denote the mass of entrants targeting protected and unprotected markets, respectively. Equation (9) states that the mass of unprotected products decreases when incumbents successfully develop an innovation at flow probability, x_I, or entrants displace existing incumbents at rate, $x_{e,1}$. Products instead become unprotected if existing patent walls depreciate, or successful entrants become new monopolists. Conversely, Equation (23) shows that protected markets loose mass whenever entrants displace existing protected incumbents or defensive patents depreciate, and gain mass when incumbents in unprotected markets develop defensive innovations.

The mass of entrants in the two types of product markets is determined in equilibrium following the laws of motion:

$$\dot{\mu}_{e,1} = -\left(x_{e,1} + x_I\right)\mu_{e,1} + x_{e,1}\mu_1 + \delta\mu_{e,\omega}, \tag{11}$$

$$\dot{\mu}_{e,\omega} = -\left(x_{e,\omega} + \delta\right)\mu_{e,\omega} + x_{e,\omega}\mu_\omega + x_I\mu_{e,1}. \tag{12}$$

Here, the first line states that the pool of entrants in unprotected markets contracts if entrants develop innovations at rate $x_{e,1}$, or incumbents make markets protected at rate x_I. The pool of entrants in markets with $\Omega = 1$ instead grows when incumbents get displaced by successful entrants and rejoin the ranks of outsiders, or previously protected markets loose their defensive walls. Equation (12) is obtained analogously. Given research intensities, clearing in the labor markets for production and research labor require:

$$L = \int_0^1 l(i)\mathrm{d}i, \tag{13}$$

$$L^{RD} = \zeta\left(\omega x_{e,\omega}\mu_{e,\omega} + x_{e,1}\mu_{e,1}\right) + \alpha_I\frac{x_I^\gamma}{\gamma}\mu_1. \tag{14}$$

A constant growth equilibrium of this economy is defined as follows.

Definition 1 (Constant Growth Equilibrium)**.** A constant growth equilibrium is a sequence of values $\{V_t(1), V_t(\omega)\}$, production workers' and inventors' wage sequences $\left\{w_t^{RD}, w_t\right\}$, and incumbents' and entrants' R&D decisions $\{x_{I,t}, x_{e,1,t}, x_{e,\omega,t}\}$ such that, given an endowment of production labor, L, L^{RD}: (i) incumbents maximize values (5) and (6), taking entrants R&D decisions as given; (ii) entrants' R&D decisions satisfy (7) and (8) taking $V_t(1)$ as given; (iii) the distribution of incumbents and entrants across protected and unprotected markets is constant, $\dot{\mu}_1 = \dot{\mu}_\omega = \dot{\mu}_{e,1} = \dot{\mu}_{e,\omega} = 0$ in Equations (9)-(12); (iv) values in each instant are determined by (5) and (6); (v) consumers maximize utility (3), choosing consumption and R&D labor optimally; (vi) labor markets clear according to (13) and (14); (vii) product markets clear, $C_t = Y_t$; and (viii) aggregate output (4) grows at the constant rate, $g \equiv \dot{Y}_t / Y_t$.

The following proposition summarizes the properties of the constant growth equilibrium.

Proposition 2 (Existence and Uniqueness of the Constant Growth Equilibrium)**.** *For any endowments of production labor L, there exists a unique constant growth equilibrium. Denoting optimal incumbents' and entrants' choices as $x_I^\star, x_{e,\omega}^\star, x_{e,1}^\star$, and the masses of incumbents and entrants across states as $\mu_1^\star, \mu_\omega^\star, \mu_{e,1}^\star, \mu_{e,\omega}^\star$, the constant growth rate of the economy is given by*

$$g = \eta \left[x_{e,\omega}^\star \mu_{e,\omega}^\star + x_{e,1}^\star \mu_{e,1}^\star + \lambda x_I^\star \mu_1^\star \right],$$

and inventors' productivity reads

$$\frac{g}{L^{RD}} = \eta \frac{x_{e,\omega}^\star \mu_{e,\omega}^\star + x_{e,1}^\star \mu_{e,1}^\star + \lambda x_I^\star \mu_1^\star}{\zeta \left(\omega x_{e,\omega}^\star \mu_{e,\omega}^\star + x_{e,1}^\star \mu_{e,1}^\star \right) + \alpha_I \frac{(x_I^\star)^\gamma}{\gamma} \mu_1^\star}.$$

The expression for inventors' productivity can be used to obtain a first intuition of the mechanism through which higher markups lead to decreased productivity in this model. First, note that a unit of total entrants' research intensity in either protected and unprotected markets produces the same growth. Indeed, by the expression for growth, it is clear that a unit increase in either $x_{e,\omega}^\star \mu_{e,\omega}^\star$ or $x_{e,1}^\star \mu_{e,1}^\star$ gives a growth of η. However, the unit labor requirement of research intensity in protected markets is larger than in unprotected markets as a result of defensive patents. This is evident from the first term at the denominator, which shows that a unit of total research in protected markets requires $\zeta\omega$ inventors, compared with just ζ in unprotected markets. These facts immediately imply that any force that pushes entrants toward protected markets will lower their inventor productivity. One such force is an increase in incumbents' research intensity that outstrips entrants', which acts to increase $\mu_{e,\omega}$, the mass of entrants active in protected markets. As I show formally below, an increase in the markup raises the value of monopolistic positions, which pushes up both incumbents' and entrants' research intensities x_I and $x_{e,\omega}$. This results in an increase in $\mu_{e,\omega}$ only if incumbents' R&D intensity is more elastic than entrants'.

The following proposition states the main result in this section: Higher markups unambiguously increase research efforts by incumbents and entrants, as well as the share of R&D labor accruing to incumbents. The proposition also states the main condition for this result to lead to a fall in overall inventors' productivity in terms of sufficient statistics. In particular, the incumbents' R&D elasticity to markups must be larger than entrants'. This condition is satisfied in the data, as the estimates in Table 5 show that incumbents increase their share of inventors following increases in markups.

Proposition 3 (Effects of Markup Increases on Innovation)**.** *Suppose that defensive research is effective, $\omega > 1$. The constant growth equilibrium features positive incumbents' research, $x_I^\star > 0$; markup increases raise both incumbents' and entrants' research effort*

$$\frac{\partial x_I^\star}{\partial \phi} > 0,\ \frac{\partial x_{e,\omega}^\star}{\partial \phi} > 0\ \frac{\partial x_{e,1}^\star}{\partial \phi} > 0;$$

and the share of R&D labor employed by incumbents increases with markup

$$\frac{\partial \frac{L_I}{L^{RD}}}{\partial \phi} = \frac{\partial \left(\alpha_I \frac{x_I^\gamma}{\gamma} \mu_1 / L^{RD} \right)}{\partial \phi} > 0.$$

Moreover, if (i) $\lambda = 0$; (ii) inventor supply is not fully inelastic; and (iii) the model parameters are such that equilibrium incumbents' research effort is more elastic than entrants',

$$\frac{\partial x_I^\star}{\partial \phi} \frac{\phi}{x_I^\star} > \frac{\partial x_{e,\omega}^\star}{\partial \phi} \frac{\phi}{x_{e,\omega}^\star},$$

it further holds

$$\frac{\partial \left(g / L^{RD} \right)}{\partial \phi} < 0.$$

That is, an increase in the markup, ϕ, lowers equilibrium inventors' productivity.

The intuition behind this proposition is that higher markups raise the value of monopolistic positions, thus raising both entrants' and incumbents' research effort. In addition, this leads to an overall increase in inventors employed by incumbents. Importantly, this holds only if defensive R&D is effective, showing the importance of this channel to reproduce the empirical findings in Table 5. However, this finding is not sufficient to generate a fall in inventor productivity. That is because the overall change in inventors' productivity can be decomposed as

$$\mathrm{d}\left(\frac{g}{L^{RD}} \right) = \mathrm{d}\left(\frac{L_e}{L^{RD}} P_e \right) + \mathrm{d}\left(\frac{L_I}{L^{RD}} P_I \right),$$

where P_e, P_I denote entrants' and incumbents' productivity, respectively. As long as incumbents' productivity is lower than entrants, a reallocation of inventors towards dominant firm leads to a decrease in inventor productivity. However, this effect can in principle be offset by an increase in entrants' productivity. The sufficient conditions in Proposition 3 ensure that this is not the case.[19] In particular, if incumbents' research effort is more elastic than entrants', an increase in markups reduces the overall share of unprotected markets, making entry overall more difficult as well as reducing entrants' productivity. This reasoning also clarifies why the condition is sufficient. Indeed, a fall in inventors' productivity occurs as long as the increase in creative-destruction growth by entrants is not high enough relative to the increase in defensive innovation by entrants. As noted above, this condition is satisfied in the data and always appear to be in numerical simulations, regardless of specific assumptions on parameter values. In particular, the assumption $\lambda = 0$ is not necessary to generate the result, but it gives clear analytical results in the proof to this proposition.

It is important to stress that the decline in productivity requires a strictly positive elasticity of inventors' supply. As I show in Lemma 5 in the Appendix, *aggregate* inventor labor demand is increasing in markups. As a result, when inventors' supply is fully rigid, wage effects need to fully offset the increase in labor demand, and by the uniqueness of the CGE, it the allocation of inventors across incumbents and entrants is fixed. Indeed, aggregate labor demand is uniquely pinned down by research intensities through the stationary distribution.

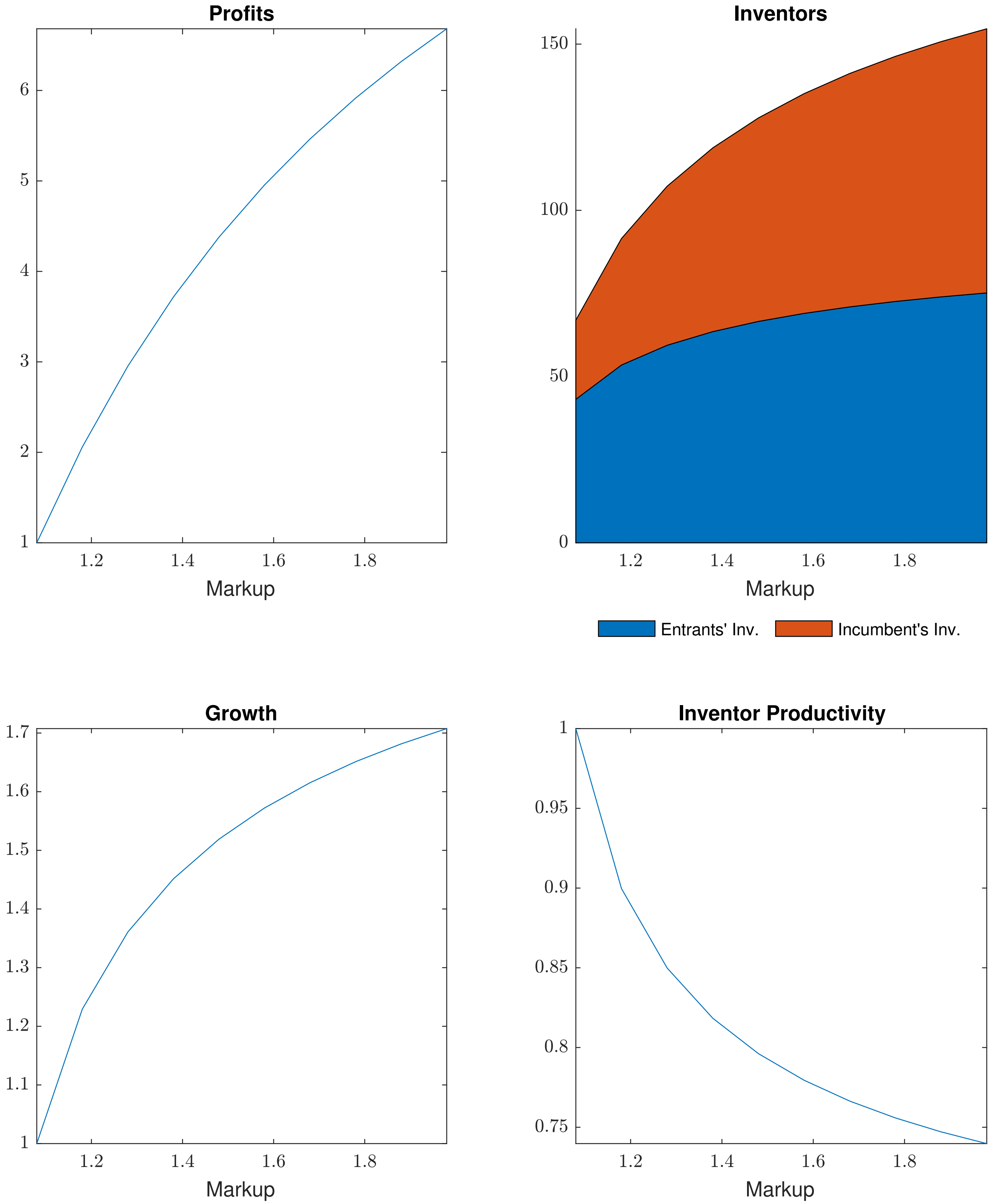
Figure 5 provides a graphical illustration of the comparative statics for an increase in markup in a calibration of the one-sector model, which follows the strategy described in detail in the following section.[20] An increase in the markup raises profits, thereby increasing the value of dominant positions and propelling both incumbents' and entrants' research. Importantly, the incumbents' inventor demand is more elastic to increased profits, which raises the overall proportion of researchers employed in defensive projects. Although overall growth increases with higher aggregate R&D, the productivity of inventors falls, as they are allocated to incumbents. Two observations are in order. First, with a fixed supply of inventors, equilibrium allocations do not change with increased profits, leaving R&D productivity unaffected. Second, in the absence of defensive innovation, the increase in profits would only result in increased entry and growth, leaving productivity constant.[21]

[19] Corollary 6 in the Appendix provides parametric sufficient conditions for the case with quadratic costs, $\gamma = 2$, which is analytically tractable.

[20] The parameters of the calibration are chosen to match moments of the R&D expenditure distribution as in Section 4.2, and the complete set of parameters is reported in Appendix.

[21] In the absence of defensive innovation, all firms operates in unprotected markets and there is no incumbent research. From the expression for growth in Proposition 3 inventor productivity is constant at η/ζ.

Figure 5: Comparative Statics for Changes in the Markup in the Single-Sector Model



Note: This figure reports the comparative statics for normalized profits, inventors, growth and inventor productivity in the single-sector model. All variables are expressed in units relative to the equilibrium with $\phi = 1.08$. The calibration follows the same strategy as in Section 4.2.2, assuming that the economy is only composed of a single sector. I set the elasticity of inventor supply at $\varphi = 1$.

4.2 Calibration and Policy

In this section, I calibrate a two-sector extension of the model presented in the previous Section, with two main objectives. First, I want to analyze misallocation *across* sectors, and show that, under a realistic calibration, this extension can qualitatively reproduce the main findings of the empirical analysis. I focus on the benchmark with *fixed* inventor supply where, as seen in the previous section, markups have no effect on inventors' productivity within sectors. This choice excludes that my findings are driven by within-sector misallocation alone. Second, I show that growth is maximized when R&D subsidies are allocated to entrants in the less competitive sector.

4.2.1 Model description

The consumption good in the economy is given by the Cobb Douglas aggregate:

$$\ln Y_t = \beta_1 \ln Y_{1,t} + \beta_2 \ln Y_{2,t}, \quad \beta_1 + \beta_2 = 1$$

where $Y_{1,t}$, $Y_{2,t}$ are produced as in Section 4.1.1, and the markup parameter, ϕ, is allowed to vary across the two sectors.

The household side of the economy is unchanged relative to the one-sector model. In this section, however, the supply of inventors is assumed to be fully rigid and given by $L^{RD} = 100$. This assumption is motivated by two considerations. First, a fixed inventor supply captures an aggregate scarcity of inventors, allowing me to focus exclusively on their allocation across sectors. Second, as discussed above, in this benchmark markup changes have no effect on misallocation if inventors are not allowed to move across sectors. Therefore, the results below isolate the effect of inventors' transitions across sectors on misallocation.

Given the presence of two markets, labor market clearing for production workers and inventors is now given by:

$$L = \sum_{i=1}^{2} \int_0^1 l_{i,j,t}\left(w^{RD}\right) \mathrm{d}j, \tag{15}$$

$$L^{RD} = \sum_{i=1}^{2} L_{i,t}^{RD}\left(w^{RD}\right), \tag{16}$$

where the subscript i denotes sectors and j the product markets in each sector. I define a constant growth equilibrium as in the previous section. Given the Cobb-Douglas assumption on the final good, growth is now given by:

$$\Delta \ln Y_t = \sum_{i=1} \beta_i g_i,$$

where g_i denotes the sector-specific output growth obtained as in Proposition 3. Appendix C.2 reports the derivations for the two-sector model and a complete description of the equations characterizing the equilibrium.

4.2.2 Calibration

I calibrate my model in order to match features of the R&D distribution and concentration around 1997, the starting year for my analysis. This approach provides conservative parameter choices to analyze the model. Under my calibration, inventor productivity falls by 2.5% over the period 1997-2012, nearly half of the 5% lower bound of my estimates. This calibration therefore produces a lower bound for the impact of optimal R&D subsidies.

Table 9b displays my choices for parameters calibrated externally. I set the discount rate to 4%, which, together with a 3% growth for my sample in 1997, implies a value for the real interest rate of 7%, in line with the long-run average before 1997. I obtain a value for β, the share of value added of each sector, from estimates of the Lerner Index in manufacturing that I obtain from the NBER-CES as described in Appendix B.5. According to these estimates, about half of the sectors (weighted by sales) for which I have data saw an increase in the Lerner Index over the period. This suggests $\beta_1 = \beta_2 = 0.5$. Since I only have the Lerner Index for about half of the sectors, I rely on the extensive literature estimating markups to set a value of $\phi = 1.08$. In particular, I follow Akcigit and Ates (2019), who calibrate the same parameter using the midpoint of estimates provided in De Loecker et al. (2020) and Eggertsson et al. (2018). As standard in the literature (see e.g., Acemoglu and Akcigit, 2012), I set the curvature of the incumbents' cost function relying on estimates by Kortum, 1993. I choose the lower bound of these estimates to minimize the asymmetry of innovation costs between incumbents and entrants, as more convex incumbents' costs mechanically make their research less effective than entrants. The rate of patent expiration comes directly from the legislative framework in the US, as established by the Uruguay Round Agreements Act of 1994. Since λ measures how radical are incumbents' innovations relative to entrants', I set $\lambda = 0.785$, the complement of the internal patent share of 21.5% estimated by Akcigit and Kerr (2018). Turning to the value of blocking patents, parametrized by ω in my model, I rely on estimates by Czarnitzki et al. (2020) and Grimpe and Hussinger (2008), who employ merger data to obtain the effect of pre-emptive patents on the value of acquired firms. Both their estimates imply an elasticity of firm's values to the share of patents with pre-emptive value of more than one. This implies that a firm with a patent portfolio composed exclusively of defensive patents is valued on average twice as much as one with only patents that have no pre-emptive value. This suggest a value of $\omega = 2$. As shown in the proof of Proposition 2, my model gives an elasticity of firms' value of at most $\omega - 1$, therefore $\omega = 2$ effectively caps this elasticity to 1. I also include R&D subsidies, modeled as a percent subsidy on inventors' wages,s, and corporate taxes applied to firms instantaneous profits, τ. I set these two parameters following the values reported by Akcigit et al. (2016).

Table 9b describes my choices for the remaining parameters, which govern the scale of R&D and the growth rate in the economy. I set the incumbents' and entrants' R&D cost scale, α_I and ζ, to match the share of inventors employed by incumbent firms in 1997 and the R&D business spending as a percent of GDP, as reported by the National Science Foundation. Intuitively, the two cost parameters jointly determine the overall R&D spending in the economy, while their relative value determines the distribution of R&D spending in equilibrium. Given the estimates for α_I and ζ, I set η to match the growth in output per worker for the sectors considered in my analysis in 1997, 3.03%. All targets are matched almost exactly.[22]

4.2.3 Comparative Statics in General Equilibrium

Figure 6 displays the comparative statics for an increase in markup in sector 2, while leaving the sector 1's markup unchanged. The graphs compares the aggregates of interest across different constant growth equilibria, and each figure reports the markup of sector 2 relative to sector 1 on the x-axis.

An increase in the relative markup of sector 2 leads to a pronounced reallocation of inventors away from sector 1. In sector 2, incoming inventors are allocated disproportionately to incumbents, who expand their share of researchers relative to entrants. This leads to a decrease in overall inventor productivity of about Computing the Lerner Index on NBER-CES data as described in Appendix B.5 reveals that the markup gap between more concentrated and less concentrated sectors has increased by about 20% over the period of interest. This implies a fall in inventors' productivity of 2.5% compared to the benchmark where the two sectors have the same competitive structure. Since the supply of inventors is fixed at $L^{RD} = 100$, this results in a 2.5% fall in GDP growth, about 0.075pp. This estimate is close to the lower bound of 0.13pp implied by my estimates in Table 8. Assuming an inelastic labor supply mutes the response of inventors' productivity to increases in the markup, so it is reasonable to expect the model in this section to understate the productivity effects of increased concentration. As discussed above, this benchmark is desirable since it shuts down productivity effects unrelated to reallocation.

Figure 7 shows the changes occurring in each sector that correspond to the aggregates reproduced in Figure 6. In this figure all variables are normalized by their value in the initial equilibrium with $\phi_1 = \phi_2 = 1.08$. Starting from the upper-left panel, increasing markups raise profits in sector 2 relative to sector 1. This shift translates into an increase in inventors in sector 2 relative to sector 1. The upper-right panel shows that incumbents' inventor demand is more elastic in response to increases in the markup, which raises top firms' share in sector 2. The increase in the equilibrium inventors' wage leads to a reallocation within sector 1 as well, where entrants gain inventors relative to incumbents. As a result of these shifts, growth increases in sector 2, and falls in sector 1, while inventor productivity follows the opposite pattern.

The reallocation of inventors within sector 1 is a feature of the model that will be important to

[22] The average percentage point deviation of moments in the model from their empirical targets is less than 10^{-6}.

Table 9: Parameter Values and Sources

(a) Parameters Calibrated Externally

Parameter Name	Symbol	Value	Source/Target
Discount rate	ρ	.04	Annual real interest rate $\approx$ 7% before 1997
Value Added Share	β	.5	Share of sectors with $\uparrow$ Lerner Index
Average Sectors' Markup	ϕ	1.08	De Loecker et al., 2020 and Eggertsson et al., 2018
Innovation Cost Curvature	γ	1/.6	Lower bound of estimates in Kortum, 1993
Intensity of Patent Expiration	δ	.05	Uruguay Round Agreements Act (1994)
Share of Implemented Innovations	λ	.785	Internal patent share of 21.5% (Akcigit and Kerr, 2018)
Value of Blocking Patents	ω	2	Czarnitzki et al., 2020; Grimpe and Hussinger, 2008
R&D subsidy	$s_I = s_e$	19%	Akcigit et al., 2016
Corporate tax rate	τ	23%	Akcigit et al., 2016

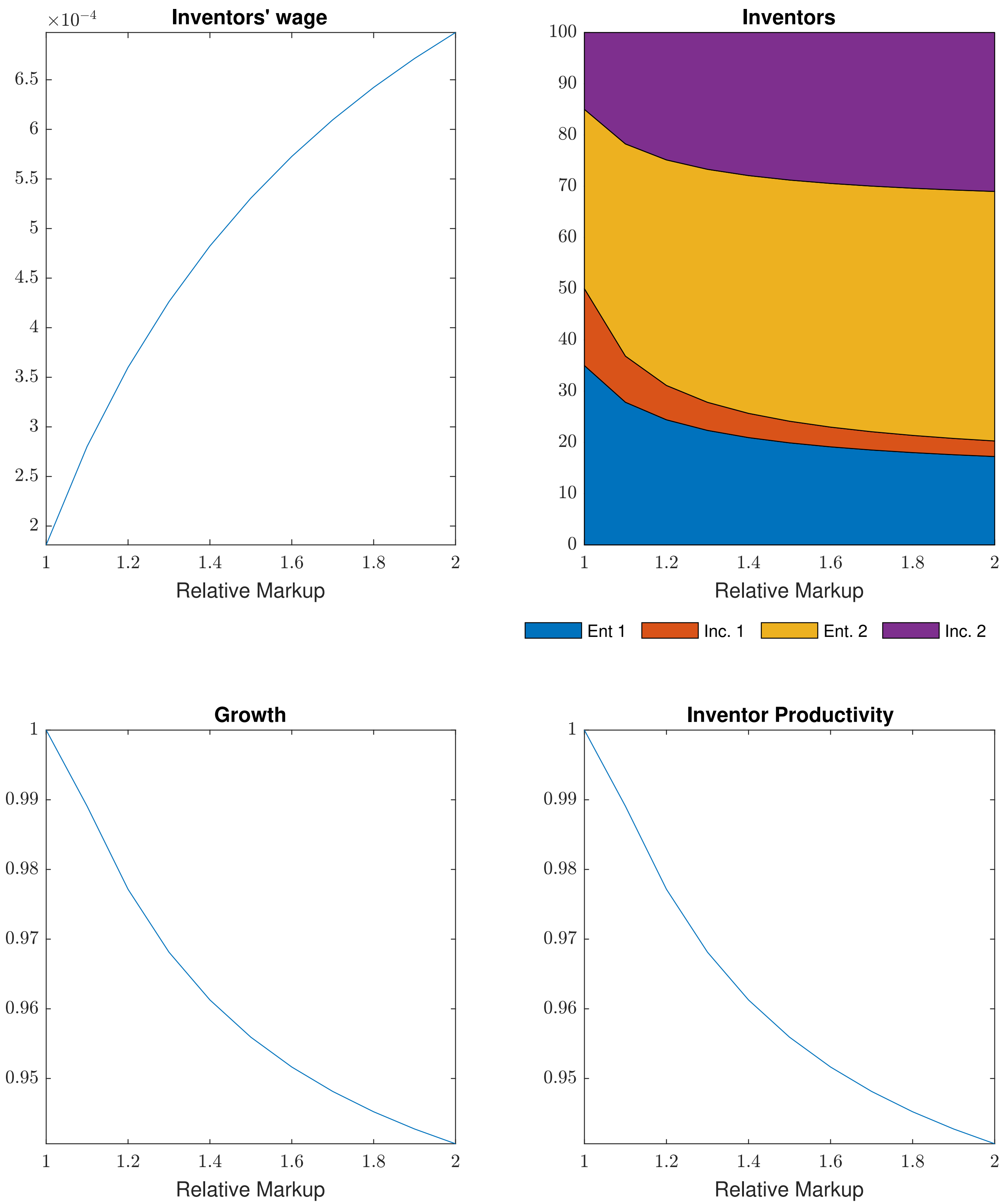
(b) Parameters Calibrated Internally

Parameter Name	Symbol	Value	Target
Incumbent Costs	α_I	21.97	Top 10% Firms' Inventor Share, 1997: 30.3%
Entrants' Costs	ζ	4.75	Business R&D Share over GDP, 1997: 1.81%
Innovation Step	η	0.0047	Output per Worker Growth, 1997: 3.03%

interpret the policy results in the following section. The contrasting behavior of sectors 1 and 2 reveals the crucial role of the overall R&D activity in shaping inventor productivity. The incentives to conduct defensive innovation increase when a larger number of inventors is employed in the sector. Therefore, while the movements of inventors from sector 1 to sector 2 are overall detrimental to growth, they are not for R&D productivity in sector 1. Since incumbents there face a lower risk of being displaced by entrants, they reduce their efforts in defensive innovation, which increases R&D productivity. However, overall growth is lower since less R&D resources are available to this sector in equilibrium. The findings in this section suggest that inventor reallocation away from competitive sectors has both costs, resulting from a reduction in sectoral growth, and benefits, coming from a more efficient distribution of resources within this sector.

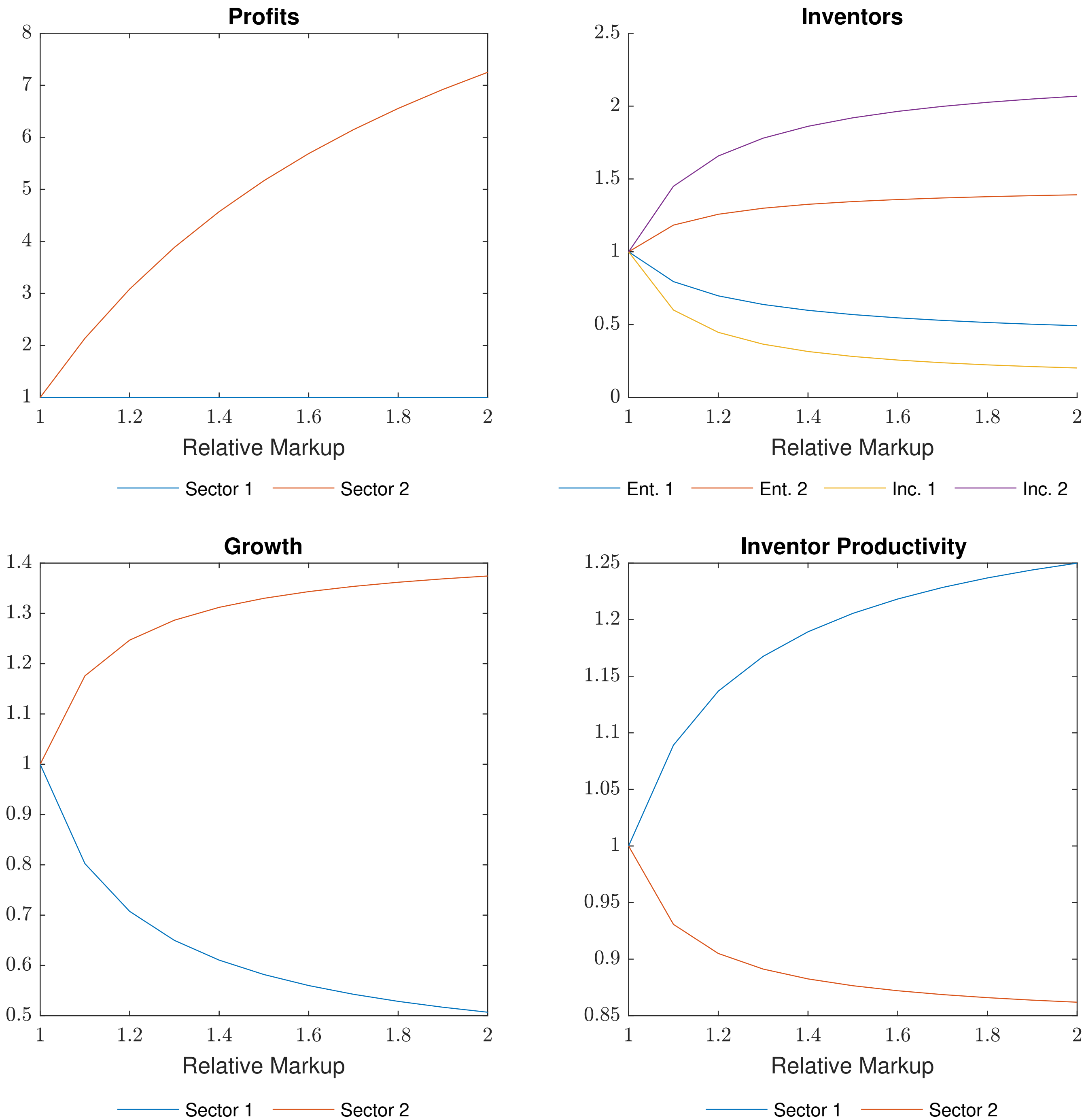
This comparative static exercise also clarifies how defensive innovation undergirds misallocation across sectors. In the absence of defensive innovation, inventors just reallocate across sectors, leaving productivity unaffected. As a result, overall growth is also unchanged, since the Cobb-Douglas assumption with $\beta = 0.5$ gives aggregate growth as the simple average of growth in the two sectors.

Figure 6: Comparative Statics in Sector 2's Markup Relative to Sector 1 in the Two-Sector Model, Economy Aggregates



Note: This figure reports the comparative statics for normalized profits, inventors, growth and inventor productivity in the two-sector model. In all figures, the x-axis reports the markup of sector 2 relative to sector 1. The parameters used to produce this figure are reported in Tables 9a and 9b.

Figure 7: Comparative Statics in Sector 2's Markup Relative to Sector 1 in the Two-Sector Model, Sector-level Aggregates



Note: This figure reports the comparative statics for normalized profits, inventors, growth and inventor productivity in the two-sector model. In all figures, the x-axis reports the markup of sector 2 relative to sector 1. The parameters used to produce this figure are reported in Tables 9a and 9b.

4.2.4 Growth-Maximizing Policy

I now turn to calculating the combination of R&D subsidies that maximizes growth. I assume that the planner wishes to maximize growth under a set of different constraint on the instruments available. In particular, I assume that the planner cannot alter the nature of innovation, that is, the planner cannot distinguish productive from unproductive projects, and cannot forbid registering patents that have no productive value. In this model, eliminating protection for incumbents would lead to a first-best where only entrants conduct R&D and reallocation could only promote growth, as discussed in the previous section.

I start from the 2012 equilibrium of my model economy, where the gap in markups between sector 2 and sector 1 is 20%, and all firms receive a 19% subsidy to inventors' wages and incur a 23% tax on profits. I then evaluate numerically three cost-neutral alternatives relative to this benchmark, that is, I constrain the planner to leave the expenditure on R&D subsidies as a percentage of GDP fixed at the 2012 benchmark. In the first scenario, the planner is allowed to distribute subsidies freely and can condition the allocation on both the state of the market (protected or unprotected) and the identity of the receiving firm (entrant or incumbent). In the other two scenarios, I only allow the planner to act on one of these dimensions at a time. That is, the planner can either control the cross-sector distribution of funds, but not the allocation across incumbents and entrants, or vice versa.

The results of this exercise are presented in Table 10. The 2012 equilibrium is reported in Columns 1 and 2. For reference, the 1997 calibrated model has two identical sectors, which share the stock of inventors equally. Within each sector incumbents have 30.3% of the overall inventors employed, and GDP growth is 3% per annum, as reported in Table 9b. In the 2012 baseline, the distribution of inventors is tilted toward the second sector, where markups have increased. This results in a fall in annual GDP growth of .07pp, about 2.5% of the 1997 benchmark. As shown in the graphs above, this new equilibrium sees a larger share of inventors allocated to incumbents in the second sector, which increases its growth relative to its more competitive counterpart. However, productivity declines because of higher defensive innovation by incumbents. Columns 3 and 4 report the optimal cost-neutral R&D subsidies chosen by a growth-maximizing planner. Somewhat surprisingly, the most efficient allocation of funds turns out to be a subsidy to entrants in the more concentrated sector only. Indeed, defensive innovation is inefficient because it makes entrants' R&D less productive. This is the main friction that the growth-maximizing planner wishes to remove. Figure 7 shows that the outflow of inventors from sector 1 increases inventor productivity, as lower R&D by entrants depresses defensive innovation by incumbents. It is therefore undesirable to reallocate inventors to entrants in sector 1, where barriers to entry are now naturally lower. Conversely, the optimal policy acts directly on the higher barriers now present in sector 2, subsidizing inventors' wages for entrants. Consistent with this argument, growth is not maximized when the planner allocates R&D subsidies to a single sector, without condition on the identity of the firm. This scenario is reported in columns 5 and 6. The planner subsidizes the more competitive sector 1, which brings annual growth up to 2.99%, recovering most of the lost ground relative to the 1997 benchmark. However, subsidies now make defensive innovation cheaper for incumbents, as well as more attractive due to increased entry. If a sector-specific subsidy to entrants is politically unfeasible, a viable alternative is a blanket entry subsidy as reported in Columns 7 and 8. The resulting growth in annual output of 3.38% exceeds the starting 1997 equilibrium.

To conclude, the policy analysis suggests that entry subsidies are the most effective policy to counter the friction introduced by defensive innovation in this model economy. The best approach is to subsidize entrants in less competitive sectors, where this friction is most pronounced, increasing growth by 0.5pp per annum. A more feasible uniform R&D subsidy to entrants produces quantitatively similar effects. Conversely, sector-specific subsidies to reallocate inventors to more competitive sectors are less effective, since incumbents use them to conduct pre-emptive innovation, precisely the source of inefficiency that the planner wishes to contrast.

5 Conclusion and Future Work

In this paper, I propose and document a novel explanation for the observed decline in growth and R&D productivity over the last few decades.

My empirical results show that increasing misallocation of inventors across different product markets can account for up to 27% of the observed decline in output per worker growth in the sectors I analyze. This misallocation stems from uneven increases in concentration across product markets that are accompanied by a larger share of inventors accruing to less competitive sectors. I interpret my findings as resulting from an increase in defensive innovation in concentrated sectors. Such R&D activities are conducted with the primary aim of blocking further entry and are reflected in a decline in patents' forward citations and an increased share of inventors employed by the largest incumbents.

The theoretical analysis examines the effects of defensive innovation in a Schumpeterian model of creative destruction, where incumbents can conduct defensive innovation to raise new entrants' costs. In the model, pre-emptive innovation is the driving force behind misallocation across sectors. Having established the importance of this mechanism, I employ a calibrated two-sector version of the model to study the growth-maximizing allocation of R&D subsidies across sectors, as well as between incumbents and entrants. My analysis suggests that R&D subsidies provided to entrants constitute the most effective policy, directly tackling the friction generated by defensive innovation and potentially increasing growth by 17% of my baseline (0.5pp in absolute terms).

Two main directions for future research stand out. The first would be to investigate the validity of my findings in an international context, given that results in the literature of competition and innovation depend strongly on the country and the period analyzed. The second would be to conduct a thorough investigation of the evolution, causes, and consequences of pre-emptive innovation.

Table 10: Comparison of R&D Policies in the Two-Sector Model

	Baseline		Optimal Cost-Neutral		Cost-Neutral Sector		Cost-Neutral Entry	
	Sector 1 (1)	Sector 2 (2)	Sector 1 (3)	Sector 2 (4)	Sector 1 (5)	Sector 2 (6)	Sector 1 (7)	Sector 2 (8)
R&D Subsidies:								
s_I	19%	19%	0%	0%	46.17%	0%	0%	0%
s_e	19%	19%	0%	41.78%	46.17%	0%	29%	29%
Aggregates:								
L_I^{RD}	6.70	24.87	6.37	15.95	10.83	19.51	4.83	18.45
L_e^{RD}	24.41	44.02	23.87	53.81	30.24	39.42	27.41	49.30
L_{TOT}^{RD}	31.11	68.89	30.25	69.75	41.07	58.93	32.25	67.75
Sector Growth	2.12%	3.74%	2.08%	4.78%	2.61%	3.36%	2.45%	4.31%
GDP Growth	2.93%		3.43%		2.99%		3.38%	

Note: The figures reported in this table give the optimal allocation of R&D subsidies and the resulting aggregate outcomes for a planner wishing to maximize aggregate growth in the economy. The column headings refer to the various scenarios described above. “Baseline” refers to the subsidy allocation reflecting the 2012 equilibrium, where subsidies do not condition on sectors or the position of firms within sectors; “Optimal Cost-Neutral” refer to the scenario where the planner is allowed to freely allocate R&D subsidies subject to the constraint that overall R&D subsidy expenditure as a percentage of GDP is held fixed at its 2012 benchmark; “Cost-Neutral Sector” consider a scenario where the planner can choose which sector to allocate funds to, but not which firms within the sector should receive the subsidy; “Cost-Neutral Entry” computes the optimal universal entry subsidy, under the assumption that the planner cannot condition its reception on the sector firms operate in.

References

D. S. Abrams, U. Akcigit, and J. Grennan. Patent Value and Citations: Creative Destruction or Strategic Disruption? Working Paper 19647, National Bureau of Economic Research, Nov. 2013.

D. Acemoglu and U. Akcigit. Intellectual Property Rights Policy, Competition and Innovation. *Journal of the European Economic Association*, 10(1):1–42, Feb. 2012.

D. Acemoglu, U. Akcigit, H. Alp, N. Bloom, and W. Kerr. Innovation, Reallocation, and Growth. *American Economic Review*, 108(11):3450–3491, Nov. 2018.

D. Acemoglu, D. Autor, and C. Patterson. Bottlenecks: Sectoral Imbalances and the US Productivity Slowdown. Working Paper, MIT, 2021.

D. Acemoglu, U. Akcigit, and M. A. Celik. Radical and Incremental Innovation: The Roles of Firms, Managers and Innovators. *American Economic Journal: Macroeconomics*, Forthcoming.

P. Aghion and P. Howitt. A Model of Growth Through Creative Destruction. *Econometrica*, 60(2):323–351, 1992.

P. Aghion, C. Harris, P. Howitt, and J. Vickers. Competition, Imitation and Growth with Step-by-Step Innovation. *The Review of Economic Studies*, 68(3):467–492, 2001.

P. Aghion, N. Bloom, R. Blundell, R. Griffith, and P. Howitt. Competition and Innovation: An Inverted-U Relationship. *The Quarterly Journal of Economics*, 120(2):701–728, 2005.

U. Akcigit and S. T. Ates. What Happened to U.S. Business Dynamism? Working Paper 25756, National Bureau of Economic Research, Apr. 2019.

U. Akcigit and S. T. Ates. Slowing Business Dynamism and Productivity Growth in the United States. Working Paper, 2020.

U. Akcigit and S. T. Ates. Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory. *American Economic Journal: Macroeconomics*, 13(1):257–298, Jan. 2021.

U. Akcigit and W. R. Kerr. Growth through Heterogeneous Innovations. *Journal of Political Economy*, 126(4):1374–1443, July 2018.

U. Akcigit, D. Hanley, and S. Stantcheva. Optimal Taxation and R&D Policies. Technical Report w22908, National Bureau of Economic Research, Cambridge, MA, Dec. 2016.

D. Argente, S. Baslandze, D. Hanley, and S. Moreira. Patents to Products: Product Innovation and Firm Dynamics. *SSRN Electronic Journal*, 2020.

D. Autor, D. Dorn, L. F. Katz, C. Patterson, C. Booth, and J. V. Reenen. The Fall of the Labor Share and the Rise of Superstar Firms. *Quarterly Journal of Economics*, Forthcoming.

P. Azoulay, I. Ganguli, and J. Graff Zivin. The mobility of elite life scientists: Professional and personal determinants. *Research Policy*, 46(3):573–590, Apr. 2017.

S. Barkai. Declining Labor and Capital Shares. *The Journal of Finance*, 75(5):2421–2463, 2020.

D. Bass and J. Brustein. Big Tech Swallows Most of the Hot AI Startups. *Bloomberg.com*, Mar. 2020.

V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, Oct. 2008.

N. Bloom, C. I. Jones, J. Van Reenen, and M. Webb. Are Ideas Getting Harder to Find? *The American Economic Review*, 110(4):41, 2020.

D. Czarnitzki, K. Hussinger, and B. Leten. How Valuable are Patent Blocking Strategies? *Review of Industrial Organization*, 56(3):409–434, May 2020.

J. De Loecker, J. Eeckhout, and G. Unger. The Rise of Market Power and the Macroeconomic Implications*. *The Quarterly Journal of Economics*, 135(2):561–644, May 2020.

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

J. G. Fernald and C. I. Jones. The Future of US Economic Growth. *American Economic Review*, 104(5): 44–49, May 2014.

N. Goldschlag, T. J. Lybbert, and N. J. Zolas. An 'Algorithmic Links with Probabilities' Crosswalk for USPC and CPC Patent Classifications with an Application Towards Industrial Technology Composition. Technical Report CES-16-15, US Census Bureau, 2016.

R. Gordon. *The Rise and Fall of American Growth: The U.S. Standard of Living since the Civil War*. Princeton University Press, Princeton, NJ, 2016.

C. Grimpe and K. Hussinger. Pre-empting technology competition through firm acquisitions. *Economics Letters*, 100(2):189–191, Aug. 2008.

G. Grullon, Y. Larkin, and R. Michaely. Are US Industries Becoming More Concentrated?*. *Review of Finance*, 23(4):697–743, July 2019.

D. Guellec, C. Martinez, and P. Zuniga. Pre-emptive patenting: Securing market exclusion and freedom of operation. *Economics of Innovation and New Technology*, 21(1):1–29, Jan. 2012.

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

B. H. Hall, A. B. Jaffe, and M. Trajtenberg. The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. Working Paper 8498, National Bureau of Economic Research, Oct. 2001.

K. Jo. Defensive Innovation and Firm Growth in the U.S.: Impact of International Trade. Working Paper, 2019.

J. Keil. The trouble with approximating industry concentration from Compustat. *Journal of Corporate Finance*, 45:467–479, Aug. 2017.

S. Kortum. Equilibrium R&D and the Patent–R&D Ratio: U.S. Evidence. *The American Economic Review*, 83(2):450–457, 1993.

C. Metz. Tech Giants Are Paying Huge Salaries for Scarce A.I. Talent. *The New York Times*, Oct. 2017.

E. Moretti and D. J. Wilson. The Effect of State Taxes on the Geographical Location of Top Earners: Evidence from Star Scientists. *American Economic Review*, 107(7):1858–1903, July 2017.

P. Shu and C. Steinwender. The Impact of Trade Liberalization on Firm Productivity and Innovation. *Innovation Policy and the Economy*, 19:39–68, Jan. 2019.

TalentSeer. 2020 AI Talent Report: Current Landscape & Market Trends. https://www.talentseer.com/2020-ai-talent-report, 2020.

A Data Construction Details

A.1 Knowledge Markets

Rescaling Inventor Flows As explained in the main text, the measure of inventor flows aims to capture the strength of the connections between two sectors. I take several steps to ensure that I do not overestimate these connections and to normalize them to account for the size of sending and receiving sectors.

As a first step, I build normalized directed flows for each inventor i in order to avoid double counting. For example, for transitions between sector 1 and 2, I define:

$$\text{flow}_{1\to 2,i,t} \equiv \frac{\sum \mathbf{1}\{i \text{ moves } 1 \to 2 \text{ in } t\}}{\sum_{j,k} \mathbf{1}\left\{i \text{ moves } j \to k \text{ in } t\right\}} \times \alpha_i.$$

This measure attributes a fraction of the effective inventor fixed effect α_i to each transition in proportion to the number of overall inventor i's transitions across sectors in each year. The first term in this formula is precisely the share of transitions from sector 1 to sector 2 relative to overall transitions between any two sectors j and k that inventor i took part in.

Second, I compute total inflows (outflows) for each NAICS 4-digit sector, summing over all years, inventors and origin (destination) sectors. For example, inflows for sector 1are defined as:

$$\text{inflow}_1 = \sum_n \sum_t \sum_i \text{flow}_{n\to 1,i,t},$$

where n denotes origin NAICS sectors, t years, and i inventor identifiers.

Third, I proceed to compute the share of directed flows between each pair of sector as a share of total inflows or outflows. For example, the share of inflows coming from sector 2 and entering sector 1 is defined as:

$$\text{share}_{1\leftarrow 2} = \frac{\sum_t \sum_i \text{flow}_{2\to 1,i,t}}{\text{inflow}_1}.$$

In this example, this measure captures the relative importance of inflows from sector 2 for the overall number of inventors received by sector 1. However, this measure can still overstate flows from large to small sectors, or vice versa. As a result, and since I need undirected flows to apply the Louvain algorithm, I define network edge weights starting from an average of the above shares of inflows and outflows for each sector and taking the minimum between the two measures as follows:

$$W_{12} = W_{21} = \min\left\{\frac{\text{share}_{1\leftarrow 2} + \text{share}_{1\rightarrow 2}}{2}, \frac{\text{share}_{1\leftarrow 2} + \text{share}_{1\rightarrow 2}}{2}\right\},$$

where $W_{12} = W_{21}$ since the final network is undirected.

Modularity Maximization Formula and Algorithm In order to identify knowledge markets from the network constructed above, I employ the Louvain community detection algorithm (Blondel et al., 2008). This algorithm maximizes the modularity of the network, Q, assigning each sector i to one of N *non-overlapping* communities c. Accordingly, the objective function for this problem is given by:

$$\max_{N} \max_{(c_1,\ldots,c_N)} Q \equiv \frac{1}{2\boldsymbol{W}} \sum_{ij} \left[W_{ij} - \frac{\boldsymbol{W}_i \boldsymbol{W}_j}{2\boldsymbol{W}} \right] \mathbf{1}\left\{c_i = c_j\right\},$$

where W_{ij}, weight of the edge connecting node i to j, and bold variables denote other summations for ease of notation. In particular, I define $\boldsymbol{W_i} \equiv \sum_k \sum_i W_{ik}$, as the sum of weights for edges with one end in node i, and the sum of all weights in the network, respectively. The indicator $\mathbf{1}\left\{c_i = c_j\right\}$ takes a value of 1 when nodes i and j belong to the same community. Note that the maximization is carried out both over the number of communities and the assignment of nodes to each community. This measure can be interpreted considering that $\frac{\boldsymbol{W}_i \boldsymbol{W}_j}{2\boldsymbol{W}}$, is the expected number of edges that arise between nodes i and j in a random network. Therefore, modularity maximizes the distance between the density of linkages within communities W_{ij} relative to the overall density of links that would arise randomly.

Since looping over all the permutations of nodes and community is numerically unfeasible, the Louvain algorithm follows an iterative procedure to maximize modularity. First, it assigns each node to its own community. Then, it repeats iteratively the following three steps:

1. Compute local deviations in modularity from reassigning the node to neighboring communities;

2. Assign nodes to communities following the local improvement granting the highest modularity increase;

3. Redefine a network with new communities as nodes.

These steps are repeated until there is no significant improvement in modularity for further steps.

B Additional Results and Robustness

B.1 Results on Overall Inventor Shares

Table 11 reports the effect of concentration increases on the share of inventors across all knowledge markets. While the correlation is positive and significant when some outliers are removed, this relation is not robust to the inclusion of all observations or the the alternative trimming procedure provided by the Mahalanobis distance. This result is unsurprising in light fo two points discussed in the main text. First, as highlighted in Section 2.2, if ordinary flows of inventors across unrelated sectors are small or absent, we should not expect any effect of changes in these sectors' characteristics on the distribution of inventors. Second, the findings reported in Table 3 suggest that cross-knowledge-market flows are not significant, as apparent from a comparison of specifications with and without knowledge-market fixed effects. The results presented in this section therefore speak to the importance of accurately delineating labor markets for inventors when assessing their flows across product markets.

Table 11: Regressions of Change in Total Inventors' Share over Change in HHI Lower Bound, Long-Difference, 1997-2012

	Ch. Total Eff. Inv. Share (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta\underline{\text{HHI}}$	0.297	1.692	1.328*	1.532*	0.271	1.889
	(2.007)	(1.956)	(0.649)	(0.696)	(2.038)	(2.023)
Δ log Sales	0.460	0.436	0.133**	0.109*	0.464	0.472
	(0.281)	(0.292)	(0.047)	(0.047)	(0.283)	(0.312)
Knowledge Market FE		✓		✓		✓
Sample	Full Sample	Full Sample	Trim Outliers	Trim Outliers	Mahalanobis 5%	Mahalanobis 5%
Weight	Sales	Sales	Sales	Sales	Sales	Sales
Observations	157	153	147	143	150	139

Note: Regressions weighted by sales in 2012; Robust standard errors in parentheses; Symbols denote significance levels $(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001)$; Checkmarks indicate the inclusion of fixed effects. Please refer to notes in Table 3 for further details.

B.2 Using the Raw Number of Inventors instead of Fixed-Effects

This Appendix reports the results for the main analysis presented in Section 3.2 using the raw number of total inventors instead of the fixed effects from regression (1), which might be inconsistently estimated. The following Tables, to be compared with Tables 2 and 3 in the main text, show that the results are qualitatively unchanged. Looking at the scale of the y-axis in panel (a) of Figure 8, it is apparent that the shares of the raw number of inventors are more volatile, and presents larger changes. This is easily explained by the fact that differences in research requirements across patent classes, firms and years are not absorbed as in the effective inventor measure. This greater variability simply results in larger and noisier coefficients, which nevertheless remain positive and significant.

Table 12: Regressions of Change in 4-digit Knowledge Market Share of Total Inventors over Change in HHI Measures, Long-Differences, 1997-2012

	Δ Inventor Share (pp) (1)	(2)	(3)	(4)	(5)	(6)
$\Delta \underline{HHI}$	74.172+		73.706+		74.177+	
	(40.957)		(41.600)		(41.047)	
Δ HHI		71.749**		71.997**		71.583**
		(24.464)		(25.060)		(24.433)
Knowledge Market FE						
Sample	Full Sample	Full Sample	Trim Outliers	Trim Outliers	Mahalanobis 5%	Mahalanobis 5%
Weight	Sales	Sales	Sales	Sales	Sales	Sales
Observations	157	80	155	79	150	72

Note: Regressions weighted by sales in 2012; Robust standard errors in parentheses; Symbols denote significance levels $(+\ p < 0.1,^{*}\ p < 0.05,^{**}\ p < .01,^{***}\ p < .001)$; Checkmarks indicate the inclusion of fixed effects. This Tables presents the results of specifications (2), when the outcome is the share of total inventors of sector p over total inventors in knowledge market k, and the independent variable is the change in the lower bound of the Herfindal-Hirschman Index for product market p, as implied by Economic Census concentration ratios, or the HHI index reported in the Economic Census. "Full Sample", "Trim Outliers" and "Mahalanobis 5%" refer to the samples described in the main text.

Table 13: Regressions of Change in 4-digit Knowledge Market Share of Total Inventors over Change in HHI Lower Bound, Long-Differences, 1997-2012

(a) Controlling for Change in Log Real Sales

	Δ Inventor Share (pp)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \underline{HHI}$	71.724+	67.160+	71.308+	67.860+	71.772+	68.398+
	(39.265)	(37.176)	(40.036)	(37.518)	(39.316)	(37.717)
Δ log Sales	1.864*	1.422*	1.688*	1.381+	1.878*	1.443+
	(0.766)	(0.717)	(0.736)	(0.711)	(0.774)	(0.745)
Knowledge Market FE		✓		✓		✓
Sample	Full Sample	Full Sample	Trim Outliers	Trim Outliers	Mahalanobis 5%	Mahalanobis 5%
Weight	Sales	Sales	Sales	Sales	Sales	Sales
Observations	157	156	155	154	150	142

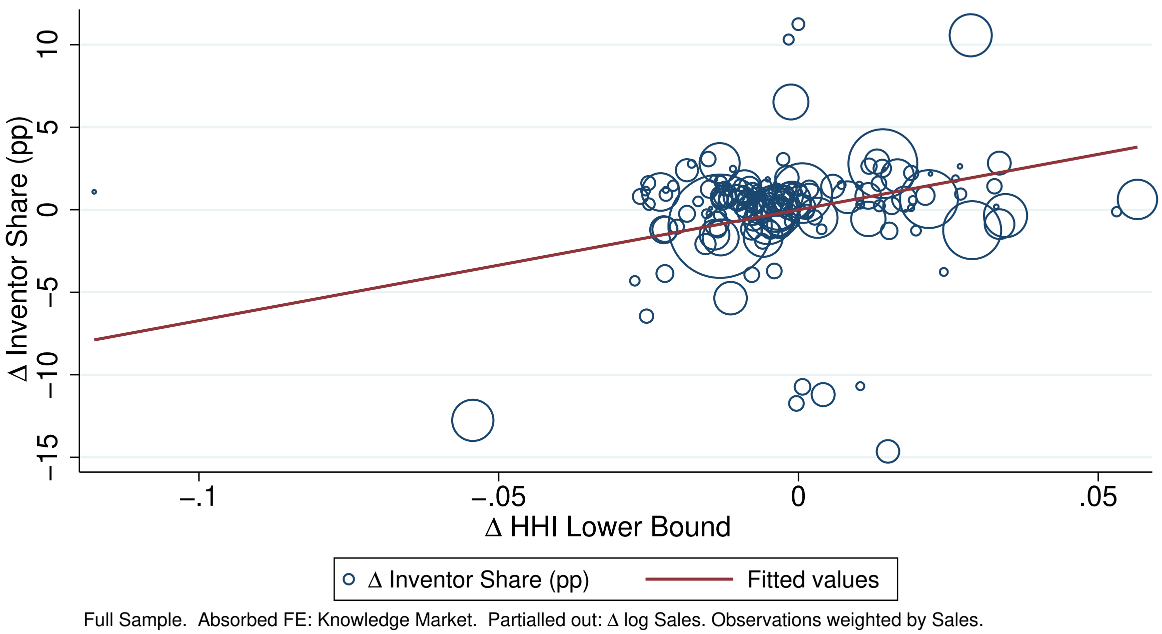
(b) Controlling for Change in Log Real Sales per Company

	Δ Inventor Share (pp)					
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \underline{HHI}$	104.562*	81.339+	103.402+	82.040+	104.355*	82.964+
	(51.534)	(43.722)	(52.824)	(43.556)	(51.356)	(46.147)
Δ log Size	0.571	-0.277	0.196	-0.515	0.571	-0.656
	(1.013)	(0.809)	(0.920)	(0.793)	(1.048)	(1.049)
Knowledge Market FE		✓		✓		✓
Sample	Full Sample	Full Sample	Trim Outliers	Trim Outliers	Mahalanobis 5%	Mahalanobis 5%
Weight	Sales	Sales	Sales	Sales	Sales	Sales
Observations	81	80	80	79	76	69

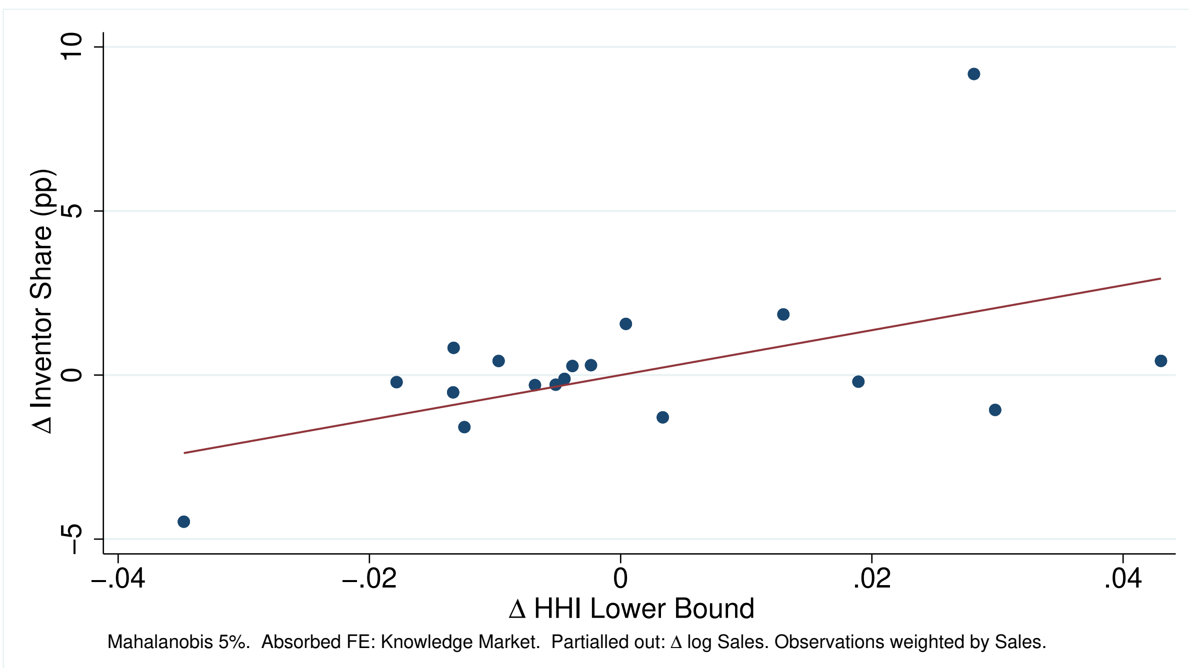
Note: Regressions weighted by sales in 2012; Robust standard errors in parentheses; Symbols denote significance levels $\left(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001\right)$; Checkmarks indicate the inclusion of fixed effects. This Tables presents the results of specifications (2) and (**??**), when the outcome is the share of effective inventors of sector p over total inventors in knowledge market k, and the independent variable is the change in the lower bound of the Herfindal-Hirschman Index for product market p, as implied by Census concentration ratios. “Full Sample”, “Trim Outliers” and “Mahalanobis 5%” refer to the samples described in the main text.

Figure 8: Residualized Scatter Plots Corresponding to Selected Columns in Table 13, Panel (a)

(a) Raw Scatter Plot, Specification in Column (2)



(b) Binned Scatter Plot, Specification in Column (6)



Note: This figure presents residualized scatter plots of the change in the share of effective inventors of sector p over total inventors in knowledge market k, over the change in the lower bound of the Herfindal-Hirschman Index for product market p, as implied by Census concentration ratios. The upper panel reports the data corresponding to the full sample, where both variables have been residualized by change in log real sales and knowledge market fixed effects. The size of the markers is proportional to the weight of each observation in the regression, corresponding to total sector sales in 2012. The regression line corresponds to the coefficient on the change in HHI lower bound reported in Column (2) of Table 13. The lower panel presents a binned scatter plot on the sample where the observations with the highest 5% Mahalanobis distance from sample centroid have been removed. Observations are aggregated using sales weights and the regression line results from the specification in Column (6) of Table 13.

B.3 Using a Quartic in Sales as Size Control

This Section displays the results of estimating the specification in Table 3 using the changes in the terms of a fourth-degree polynomial in sales rather than log-sales. This flexible control specification ensures that my main findings do not rely on the specific functional form that I assumed above. Table 14 reports the result of this exercise using both effective inventors (Columns (1) and (2)) and raw inventor counts (Columns (3) and (4)) to compute sector shares. Recall that when using raw inventor counts, knowledge markets are also constructed according to this measure. As clear from a comparison of Columns (1) with (2), and (3) with (4), these two specifications produce statistically undistinguishable results.

Table 14: Regressions of Change in 4-digit Knowledge Market Share of Inventors over Change in HHI Lower Bound, Long-Differences, 1997-2012

	Δ Inventor Share (pp)			
	(1)	(2)	(3)	(4)
Δ$\underline{\text{HHI}}$	22.509*	24.083*	67.160+	74.769+
	(10.848)	(10.565)	(37.176)	(39.225)
Δ log Sales	0.548*		1.422*	
	(0.243)		(0.717)	
Δ Sales ($ bn)		2.617*		6.382+
		(1.108)		(3.365)
ΔSales^2		-0.749		-1.749
		(0.482)		(1.468)
ΔSales^3		0.081		0.165
		(0.076)		(0.232)
ΔSales^4		-0.003		-0.005
		(0.003)		(0.009)
4D Knowledge Market FE	✓	✓	✓	✓
Sample	Full Sample	Full Sample	Full Sample	Full Sample
Weight	Sales	Sales	Sales	Sales
Observations	153	153	156	156

Note: Regressions weighted by sales in 2012; Robust standard errors in parentheses; Symbols denote significance levels $(+\ p < 0.1,^{*}\ p < 0.05,^{**}\ p < .01,^{***}\ p < .001)$; Checkmarks indicate the inclusion of fixed effects. This Tables presents the results of specifications (2), when the outcome is the share of effective inventors of sector p over total inventors in knowledge market k, and the independent variable is the change in the lower bound of the Herfindal-Hirschman Index for product market p, as implied by Census concentration ratios. "Full Sample" refers to the sample described in the main text.

B.4 Using the HHI as the Independent Variable in Patent Quality Regressions

This Section reports the results analogous to Tables 5 and 6 in the main text, using the HHI lower bound as the independent variable rather than the share of inventors. As should be expected from the high correlation between the two variables, the results are qualitatively similar to those reported there, with some distinctions. First, Table 15 confirms that there has a been a large and significant increase in the concentration of inventors among top firms in product market with increasing HHI. Differently from the main specification, this seems to be driven primarily by a fall in the share of the bottom half innovative firms. However, this does not counter the interpretation provided in the main text that increased concentration tends to reduce entry, which manifests in these regressions through a fall in the share of inventors employed by smaller, and presumably younger, firms. Table 16 shows the robustness of my findings on forward citations to the use of the HHI as well as weighting the regressions by 2012 sales, although the generality coefficient appears non-significant, as discussed in the main text.

B.5 Using the Lerner Index instead of the HHI

Following Grullon et al. (2019), I build the Lerner Index from NBER-CES data for the period 1997-2012 as the ratio:

$$\text{Lerner}_{jt} = \frac{\text{vship}_{jt} - \text{pay}_{jt} - \text{matcost}_{jt} - \text{energy}_{jt}}{\text{vship}_{jt}}, \tag{17}$$

where "vship" is the total value of shipments, "pay" denotes total payrolls, "matcost" and "energy" material and energy costs, respectively, and j denotes a 6- or 4-digit NAICS sector. I build two alternative measures, one using 6-digit NAICS sectors, the original identifier in NBER-CES, and then averaging by sales at the level of 4-digit NAICS, or first aggregating the revenue and cost statistics at the level of 4-digit NAICS. Table 17, shows that the Lerner Index thus constructed is strongly correlated with the HHI measure used in the main analysis. However, the correlation is far from perfect, as suggested by the R^2, suggesting that this estimate of the Lerner Index might be excessively imprecise. Indeed, Table 18 shows that, when using this measure instead of the HHI in the main analysis, the coefficients for the regression of inventors' shares on changes in concentration stay positive, but become smaller and noisier. This suggests the potential presence of attenuation bias, a valid concern due to the fact that the above measure, not based on any structural estimation, can only imperfectly capture markups. Note that this is also due to the fact that the Lerner Index is available only for the manufacturing sectors, which make up about 60% of the sample, so its use lead to dropping a substantial amount of observations. When using fitted values from the regression in Table 17 to extend the measure to more sectors, as well as reducing the volatility of the series for available sectors, the coefficients recover magnitudes and significance close to the baseline presented in 3.

Table 15: Regressions of Change in Inventor Distribution Measures over Change in 4-digit Knowledge Market Share, Long-Difference, 1997-2012

	Δ 90/50 Quantile Ratio	Δ Top 10%/Bottom 50%	Δ Top-50/Bottom-50 Share Ratio	Δ Top 10%	Δ Bottom 50%
	(1)	(2)	(3)	(4)	(5)
ΔHHI	15.426*	1.793	10.566	-0.085	-0.409*
	(6.848)	(5.797)	(8.078)	(0.539)	(0.188)
Δ log Sales	0.048	0.464	0.340	0.036	-0.000
	(0.154)	(0.349)	(0.407)	(0.022)	(0.008)
4D Knowledge Market FE	✓	✓	✓	✓	✓
Sample	Full Sample	Full Sample	Full Sample	Full Sample	Full Sample
Weight	Sales	Sales	Sales	Sales	Sales
Observations	118	118	118	118	118

Note: Regressions weighted by sales in 2012; Robust standard errors in parentheses; Symbols denote significance levels $\left(+\ p < 0.1,^{*}\ p < 0.05,^{**}\ p < .01,^{***}\ p < .001\right)$; Checkmarks indicate the inclusion of fixed effects. Please refer to notes in Table 3 for further details. Column (1) uses the ratio in the 90 percentile of effective inventors to the median as the outcome variable. Columns (2) and (3) instead present the share ratio, that is the share of effective inventors accruing to the top 10 or 50% relative to the share accruing to the bottom 50% of the distribution within each NAICS sector.

Table 16: Regressions of Changes in Forward Citation over HHI Changes, Long-Differences, 1997-2012

(a) Full sample

	Δ log Citations/Patent (CPC)	Δ log Citations/Patent (Total)	Δ Patent Generality
	(1)	(2)	(3)
$\Delta\underline{\text{HHI}}$	-11.133**	-12.524**	-0.335
	(3.730)	(4.324)	(0.431)
Δ log Sales	-0.454*	-0.523*	-0.019
	(0.201)	(0.257)	(0.022)
Knowledge Market FE	✓	✓	✓
Sample	Full Sample	Full Sample	Full Sample
Weight	Sales	Sales	Sales
Observations	153	153	153

(b) Full sample, restricting to the middle range of the change in inventor shares (−2% to +2%)

	Δ log Citations/Patent (CPC)	Δ log Citations/Patent (Total)	Δ Patent Generality
	(1)	(2)	(3)
$\Delta\underline{\text{HHI}}$	-10.646**	-13.052**	-0.624
	(4.018)	(4.979)	(0.473)
Δ log Sales	-0.467*	-0.554*	-0.022
	(0.214)	(0.273)	(0.023)
Knowledge Market FE	✓	✓	✓
Sample	Full Sample	Full Sample	Full Sample
Weight	Sales	Sales	Sales
Observations	144	144	144

Note: Unweighted regressions; Robust standard errors in parentheses; Symbols denote significance levels $\left(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001\right)$; Checkmarks indicate the inclusion of fixed effects. This Tables presents the results of specification (**??**), when the outcome is the log-change in forward citations and the change in patent generality in sector p over the change in the share of inventors employed in sector p. Column (1) and (2) presents the results when forward citations are extrapolated the procedure Hall et al. (2000) to avoid truncation bias. A specific cite-lag distribution over 35 years is estimated for each pair of cited and citing CPC2-codes. Column (1) employs the extrapolation scheme by each pair of CPC2 cited and citing sector. Column (2) applies the extrapolation scheme to total citations received by each cited patent. Column (3) presents results on the patent generality measures. All columns exclude self-citations. Upper panel: full sample; Bottom panel: excluding sectors with absolute increase in the inventor share above 2%.

Table 17: Regressions of Changes in the Lerner Index over Changes in the HHI Lower Bound, Long-Difference, 1997-2012

	Δ Lerner Index
	(2)
$\Delta \underline{HHI}$	1.652***
	(0.257)
Observations	258
R-squared	.14

Note: Robust standard errors in parentheses; Symbols denote significance levels $\left(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001\right)$. "6d Lerner Index" refers to the Lerner Index constructed as in (17)on NAICS 6-digits averaged at the 4-digit NAICS level weighting by the value of shipments; "4d Lerner Index" is computed using 4-digit aggregates for the value of shipments, payroll and costs, summing over the NAICS 6-digit composing each sector.

Table 18: Regressions of Changes in Inventors' Share over Changes in Actual and Fitted Lerner Index, Long-Difference, 1997-2012

	Δ Inventor Share (pp)	
	(1)	(2)
Δ Lerner	0.556	
	(5.465)	
Δ Lerner (Fitted)		26.736*
		(13.363)
Knowledge Market FE		
Sample	Full Sample	Full Sample
Weight	Sales	Sales
Observations	81	157

Note: Robust standard errors in parentheses; Symbols denote significance levels $\left(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001\right)$; Observations weighted by sales. The markup change 1997-2012 is the long-difference of the Lerner Index described above. "Fitted Lerner change" is the fitted value for the Lerner index based on the estimates in 17, and extended to all available sectors in the main sample.

C Omitted Proofs and Derivations

C.1 One-sector model

Proof of Proposition 3. This proof consists of several parts. First, I show that given labor supplies, output, values and wages grow at the same constant rate, so the problem can be solved in a steady state of a normalized model. Second, show that normalized values, $v(\Omega) \equiv V_t(\Omega)/Y_t$, are uniquely determined, which gives unique research intensities and stationary distribution. Third, I derive the stationary distribution and the expression for growth and inventors' productivity. In what follows I suppress stars to denote equilibrium quantities for ease of notation.

Given an endowment, L, production labor market clearing in each period requires:

$$\int_0^1 l_{i,t}(w)d(i) = L.$$

That is,

$$L = \int_0^1 \frac{c_{i,t}}{\phi} y_{i,t}(w_t)\, d(i) = \frac{1}{\phi}\frac{Y_t}{w_t},$$

where the second equality comes from using the demand for output of product i for $y_{i,t}(w_t)$. This expression immediately implies that if Y_t grows at a constant rate, so does w_t. Labor market clearing for R&D workers reads:

$$L^{RD} = \zeta \omega x_{e,\omega}\left(\mu_{e,\omega} + \mu_{e,1}\right) + \alpha_I \frac{(x_I)^\gamma}{\gamma}\mu_1.$$

In a constant growth equilibrium (CGE), the distribution is stationary, and since the left hand side is constant, research intensities are also fixed. A contradiction arises otherwise, since the distribution is stationary only if research intensities are fixed by the LOM (46)-(49). Further, R&D labor cannot grow since the growth rate in the economy increases in total R&D labor for any given distribution, as it will be clear below. The fact that research intensities are constant immediately implies, from the optimality of $x_{e,\omega}$, that $V_t(1)$ and w_t^{RD} grow at the same rate. Indeed, from the FOC for entrants' research:

$$0 = \mathrm{d}\log x_{e,\omega,t} = \mathrm{d}\log V_t(1) - \mathrm{d}\log w_t^{RD}.$$

This result in turn implies, combined with the FOC for x_I, that $V_t(\omega)$ also grows at the same constant rate. Now consider the budget constraint of the representative household, combined with product market clearing, $Y_t = C_t$:

$$r_t A_t - \dot{A}_t + w_t^{RD} L^{RD} + w_t L = Y_t,$$

where A_t denote the household's assets, that is all firms in the economy. Therefore the above reads:

$$r_t\left(\mu_1 V_t(1) + \mu_\omega V_t(\omega)\right) - \mu_1 \dot{V}_t(1) - \mu_\omega \dot{V}_t(\omega) + w_t^{RD} L^{RD} + w_t L = Y_t$$

Dividing both sides by $V(1)$, using the Euler equation and rearranging we obtain:

$$(g+\rho)\left(\mu_1 + \mu_\omega \frac{V_t(\omega)}{V_t(1)}\right) - \mu_1 g_{V_1} - \mu_\omega \frac{V_t(\omega)}{V_t(1)} g_{V_1} + \frac{w_t^{RD}}{V_t(1)} L^{RD} = \frac{Y_t}{V_t(1)} - \frac{w_t}{V_t(1)} L.$$

By what shown above, all terms on the left hand side are constant in t, since research wages and values grows at the same rate and the distribution is stationary. Since Y_t and w_t grow at the same rate positive rate, it must be that $V_t(1)$ also grows at the same rate as Y_t. This proves that $g_{V_1} = g = g_c = g_w = g_{w^{RD}}$.

As a result, in a CGE, it is possible to define normalized constant values, $v(\Omega) \equiv V_t(\Omega)/Y_t$. The system of equations defining the recursive problem in this equilibrium reads:

$$\rho v(1) = \max_{x_I} \left\{ \left(\frac{\phi - 1}{\phi} \right) - \alpha_I \frac{x_I^{\gamma}}{\gamma} w^{RD} + x_I \left(v(\omega) - v(1) \right) - x_{e,1} v(1) \right\}, \tag{18}$$

$$\rho v(\omega) = \left(\frac{\phi - 1}{\phi} \right) + \delta \left(v(1) - v(\omega) \right) - x_{e,\omega} v(\omega), \tag{19}$$

where the left hand side comes from using the Euler equation:

$$r = g + \rho$$

Which gives

$$r \frac{V_t(\Omega)}{Y_t} - \frac{\dot{V}_t(\Omega)}{Y_t} \frac{Y_t}{\dot{Y}_t} \frac{\dot{Y}_t}{V_t(\Omega)} \frac{V_t(\Omega)}{Y_t} = \left(\rho + g \right) v(\Omega) - g v(\Omega) = \rho v(\Omega).$$

I now move to show that normalized values (18) and (19) are uniquely determined. Given entrants' decisions, and a wage rate w^{RD}, the incumbent's choice of R&D satisfies:

$$x_I = \mathbf{1} \left\{ v(\omega) - v(1) > 0 \right\} \left(\frac{v(\omega) - v(1)}{\alpha_I w^{RD}} \right)^{\frac{1}{\gamma - 1}}.$$

Entrants taking x_I as given optimally set:

$$x_{e,1} = \mathbf{1} \left\{ v(1) > 0 \right\} \frac{v(1)}{\zeta w^{RD}}, \quad x_{e,\omega} = \mathbf{1} \left\{ v(1) > 0 \right\} \frac{v(1)}{\zeta \omega w^{RD}}.$$

Note that these solutions immediately imply that the normalized value, $v(1)$, is strictly positive. Indeed, $v(1) < 0$ would imply:

$$\rho v(1) = \pi + \mathbf{1} \left\{ v(\omega) - v(1) > 0 \right\} \left(\frac{\gamma - 1}{\gamma} \left(\frac{v(\omega) - v(1)}{\alpha_I w^{RD}} \right)^{\frac{1}{\gamma - 1}} \right) \left(v(\omega) - v(1) \right)$$

where the right hand side is strictly positive. Plugging optimal solutions into the system of equations determining the value functions (42) and (43) gives:

$$\rho v(1) - \pi - \mathbf{1} \left\{ v(\omega) - v(1) > 0 \right\} \left(\frac{\gamma - 1}{\gamma} \left(\frac{v(\omega) - v(1)}{\alpha_I w^{RD}} \right)^{\frac{1}{\gamma - 1}} \right) \left(v(\omega) - v(1) \right) + \frac{v(1)^2}{\zeta w^{RD}} = 0 \tag{20}$$

$$\rho v(\omega) - \pi - \delta \left(v(1) - v(\omega) \right) + \frac{v(1)}{\zeta w^{RD} \omega} v(\omega) = 0. \tag{21}$$

The second equation gives $v(\omega)$ as the following function of $v(1)$:

$$v(\omega) = \frac{\pi + \delta v(1)}{\rho + \delta + \frac{v(1)}{\zeta w^{RD} \omega}}.$$

Suppose first that $v(\omega) < v(1)$. In this case, the first equation gives:

$$\rho v(1) + \frac{v(1)^2}{\zeta w^{RD}} - \pi = 0.$$

The roots of this equation are:

$$v_{1,2} = \frac{-\rho \pm \sqrt{\rho^2 + 4\frac{\pi}{\zeta w^{RD}}}}{\frac{2}{\zeta w^{RD}}}.$$

Since the term under the root is strictly positive, only one of these roots is admissible, so the above system is solved for a unique pair $v(1), v(\omega)$. Consider now the case $v(\omega) > v(1)$. It is straightforward to note that $v(\omega) - v(1)$ is decreasing in $v(1)$. This implies that, when rewriting (20) as

$$-\left(\frac{\gamma-1}{\gamma}\left(\frac{v(\omega)-v(1)}{\alpha_I w^{RD}}\right)^{\frac{1}{\gamma-1}}\right)(v(\omega)-v(1)) = \pi - \rho v(1) - \frac{v^2(1)}{\zeta w^{RD}}, \tag{22}$$

the left hand side is monotonically increasing in $v(1)$, while the right hand side is monotonically decreasing in $v(1)$. Further, at $v(1) = 0$, the left hand side is strictly negative, while the right hand side equals π, while for $v(1) \to \infty$, the right hand side tends to $+\infty$ while the left hand side decreases towards $-\infty$. As a result, (22) has a unique positive solution.

The uniqueness of $v(1)$ immediately implies unique $v(\omega)$ and R&D choices. Given these R&D choices, the stationary distribution satisfies

$$0 = -\left(x_I + x_{e,1}\right)\mu_1 + \delta\mu_\omega + x_{e,\omega}\mu_{e,\omega} + x_{e,1}\mu_{e,1}, \tag{23}$$

$$0 = -\left(x_{e,\omega} + \delta\right)\mu_\omega + x_I\mu_1, \tag{24}$$

$$0 = -\left(x_{e,1} + x_I\right)\mu_{e,1} + x_{e,1}\mu_1 + \delta\mu_{e,\omega}, \tag{25}$$

$$0 = -\left(x_{e,\omega} + \delta\right)\mu_{e,\omega} + x_{e,\omega}\mu_\omega + x_I\mu_{e,1}. \tag{26}$$

By equation (24):

$$x_I\mu_1 = \left(x_{e,\omega} + \delta\right)\mu_\omega$$

Since $\mu_1 = 1 - \mu_\omega$, the stationary distribution has:

$$\mu_\omega = \frac{x_I}{x_I + x_{e,\omega} + \delta},$$

$$\mu_1 = \frac{x_{e,\omega} + \delta}{x_I + x_{e,\omega} + \delta},$$

$$\begin{bmatrix} -\delta & x_{e,1} + x_I \\ x_{e,\omega} + \delta & -x_I \end{bmatrix} \begin{bmatrix} \mu_{e,\omega} \\ \mu_{e,1} \end{bmatrix} = \begin{bmatrix} x_{e,1}\mu_1 \\ x_{e,\omega}\mu_\omega \end{bmatrix}. \tag{27}$$

Since the matrix in (27) is nonsingular, $\mu_{e,\omega}$ and $\mu_{e,1}$ are uniquely determined as:

$$\mu_{e,\omega} = \frac{x_I x_{e,1}\mu_1 + \left(x_{e,1} + x_I\right) x_{e,\omega}\mu_\omega}{x_{e,\omega}\left(x_{e,1} + x_I\right) + \delta x_{e,1}},$$
$$\mu_{e,1} = \frac{\left(x_{e,\omega} + \delta\right) x_{e,1}\mu_1 + \delta x_{e,\omega}\mu_\omega}{x_{e,1}\left(x_{e,\omega} + \delta\right) + x_{e,\omega} x_I}$$

By the optimal solution for entrants:

$$x_{e,1} = \omega x_{e,\omega},$$

so (27) is solved for:

$$\mu_{e,\omega} = \frac{\omega x_I \mu_1 + \left(\omega x_{e,\omega} + x_I\right)\mu_\omega}{\omega\left(x_{e,\omega} + \delta\right) + x_I}, \tag{28}$$
$$\mu_{e,1} = \frac{\omega\left(x_{e,\omega} + \delta\right)\mu_1 + \delta\mu_\omega}{\omega\left(x_{e,\omega} + \delta\right) + x_I}. \tag{29}$$

Thus, the stationary distribution is unique.

It remains to show that equilibrium R&D labor is also unique. To show this, I prove that R&D labor demand is monotonically decreasing in wages and has:

$$\lim_{w^{RD}\to\infty} L^{RD}(w^{RD}) \le 0, \qquad \lim_{w^{RD}\to 0} L^{RD}(w^{RD}) = \infty.$$

Since the converse holds for R&D labor supply is monotonically increasing in wages and ranges between 0 and $+\infty$, this gives a unique intersection of the two schedules. First note that, if labor supply is inelastic, $\phi = 0$, equilibrium R&D labor is constant by definition. Lemma 5 below builds on this observation as well as 4 to prove that research labor demand is indeed monotonically decreasing in the wage.

Lemma 4. *Consider a steady state of the normalized one-sector model, and assume that defensive innovation is effective, $\omega > 1$. Then, $\omega v(1) > v(\omega) > v(1)$. Around a steady state, and for a fixed wage rate, w^{RD}, the normalized values, $v(1), v(\omega)$, are increasing in the markup, ϕ, and*

$$\frac{\partial v(\omega)}{\partial \phi} > \frac{\partial v(1)}{\partial \phi} > 0.$$

Proof of Lemma 4. Subtracting side by side Equation (20) from (21) gives:

$$\left(\rho + \delta + \mathbf{1}\{v(\omega) - v(1) > 0\}\left(\frac{\gamma - 1}{\gamma}\left(\frac{v(\omega) - v(1)}{\alpha_I w^{RD}}\right)^{\frac{1}{\gamma-1}}\right)\right)(v(\omega) - v(1)) = \frac{v(1)}{\zeta w^{RD}}\left(v(1) - \frac{v(\omega)}{\omega}\right)$$

Suppose that $v(\omega) < v(1)$. This implies that the left hand side of the above expression is strictly smaller than 0, while $\omega v(1) > v(1) > v(\omega)$, so the right hand side is strictly positive under the assumption $\omega > 1$. Therefore, it must be that $v(\omega) > v(1)$. If this is the case, the left hand side is strictly positive, and to avoid a contradiction it must be $\omega v(1) > v(\omega)$. Thus, $\omega v(1) > v(\omega) > v(1)$, proving the first part of the statement.

Since π is a monotone increasing function of ϕ, I prove the statement for value derivatives with respect to π. Total differentiation of the system of Equations (20) and (21) with respect to π around a CGE gives

$$\underbrace{\begin{bmatrix} \rho + \left(\frac{v(\omega)-v(1)}{\alpha_I w^{RD}}\right)^{\frac{1}{\gamma-1}} + 2\frac{v(1)}{\zeta} & -\left(\frac{v(\omega)-v(1)}{\alpha_I w^{RD}}\right)^{\frac{1}{\gamma-1}} \\ -\delta + \frac{v(\omega)}{\zeta w^{RD}\omega} & \rho + \delta + \frac{v(1)}{\zeta w^{RD}\omega} \end{bmatrix}}_{\equiv J} \begin{bmatrix} \mathrm{d}v(1) \\ \mathrm{d}v(\omega) \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathrm{d}\pi = 0. \tag{30}$$

The determinant of the Jacobian is:

$$\det J = \left(\rho + x_I + 2\omega x_{e,\omega}\right)\left(\rho + \delta + \frac{v(1)}{\zeta\omega}\right) + x_I\left(x_{e,\omega} - \delta\right) > 0.$$

Solving (30) gives:

$$\begin{bmatrix} \frac{\mathrm{d}v(1)}{\mathrm{d}\pi} \\ \frac{\mathrm{d}v(\omega)}{\mathrm{d}\pi} \end{bmatrix} = \frac{1}{\det J} \begin{bmatrix} \frac{v(1)}{\zeta w^{RD}\omega} + \rho + \delta & \left(\frac{v(\omega)-v(1)}{\alpha_I w^{RD}}\right)^{\frac{1}{\gamma-1}} \\ \delta - \frac{v(\omega)}{\zeta w^{RD}\omega} & \rho + \left(\frac{v(\omega)-v(1)}{\alpha_I w^{RD}}\right)^{\frac{1}{\gamma-1}} + 2\frac{v(1)}{\zeta w^{RD}} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Since the first row is strictly positive,

$$\frac{\mathrm{d}v(1)}{\mathrm{d}\pi} > 0.$$

Subtracting line by line gives:

$$\begin{aligned} \frac{\mathrm{d}v(\omega)}{\mathrm{d}\pi} - \frac{\mathrm{d}v(1)}{\mathrm{d}\pi} &= \frac{1}{\det J}\left[-\frac{v(\omega)}{\zeta w^{RD}\omega} - \rho + \frac{v(1)}{\zeta w^{RD}\omega} + \rho + 2\frac{v(1)}{\zeta w^{RD}}\right] \\ &= \frac{1}{\det J}\left[-\frac{v(\omega)}{\zeta w^{RD}\omega} - \frac{v(1)}{\zeta w^{RD}\omega} + 2\frac{v(1)}{\zeta w^{RD}}\right] \\ &= \frac{1}{\det J}\left[\frac{2\omega v(1) - (v(\omega) + v(1))}{\zeta w^{RD}\omega}\right] > 0 \end{aligned} \tag{31}$$

since $\omega > 1$ and $\omega v(1) > v(\omega)$, from what shown above. It follows that:

$$\frac{\mathrm{d}v(\omega)}{\mathrm{d}\pi} > \frac{\mathrm{d}v(1)}{\mathrm{d}\pi} > 0.$$

□

Lemma 5. *R&D labor demand is monotonically decreasing in the wage rate w_t^{RD}/Y_t,and:*

$$\lim_{w^{RD}\to\infty} L^{RD}(w^{RD}) \leq 0, \qquad \lim_{w^{RD}\to 0} L^{RD}(w^{RD}) = \infty.$$

Proof. Consider the equilibrium with inelastic R&D labor. By the resource constraint in the economy, it holds:

$$\rho\left(\mu_1 v(1) + \mu_\omega v(\omega)\right) + w^{RD}L^{RD} + wL = 1,$$
$$L^{RD} = \frac{\pi}{w^{RD}} - \rho\left(\mu_1\frac{v(1)}{w^{RD}} + \mu_\omega\frac{v(\omega)}{w^{RD}}\right).$$

Since the labor supply is fixed, shifts in the right hand side of this equation identify the elasticity of labor supply to various parameters. Now consider an increase in π to $\pi' > \pi$. In this case, the unique equilibrium requires:

$$\frac{\pi'}{w^{RD\prime}} = \frac{\pi}{w^{RD}}.$$

Indeed, guess that the equilibrium involves no changes in research intensities, and therefore in the stationary distribution. Then:

$$x'_{e,\omega} = x_{e,\omega} \Rightarrow \frac{v'(1)}{\zeta\omega w'^{RD}} = \frac{v(1)}{\zeta\omega w^{RD}},$$

and

$$x'_I = \left(\frac{v'(1)-v'(\omega)}{\alpha_I w'^{RD}}\right)^{\frac{1}{\gamma-1}} = \left(\frac{v(1)-v(\omega)}{\alpha_I w^{RD}}\right)^{\frac{1}{1-\gamma}} = x_I.$$

As a result:

$$\frac{v'(\omega)}{w'^{RD}} = \frac{v(\omega)}{w^{RD}}.$$

Using the expression for $v(\omega)$, and using the fact that the ratio between values and wages is the same in both equilibria, gives:

$$\frac{\pi'}{w'^{RD}} = \frac{\pi}{w^{RD}}.$$

This also ensures that:

$$\rho\frac{v(1)}{w^{RD}} = \rho\frac{v'(1)}{w'^{RD}},$$

as is easily verified plugging the above expression into (18) evaluated at $\left(v(1), w^{RD}\right)$ and $\left(v'(1), w^{RD\prime}\right)$.It remains to show that goods' market clearing holds. Before a markup change we have (in normalized values):

$$\rho\left(\mu_1 v(1) + \mu_\omega v(\omega)\right) + w^{RD}L^{RD} + wL = 1,$$
$$\rho\left(\mu_1\frac{v(1)}{w^{RD}} + \mu_\omega\frac{v(\omega)}{w^{RD}}\right) + L^{RD} = \frac{1-wL}{w^{RD}},$$

By what shown above, with an inelastic labor research labor supply, the left hand side has the same value before and after the change in instantaneous profits. Further, the linear production function implies that:

$$wL = \frac{1}{\phi},$$

therefore the right hand side can be written as:

$$\frac{\pi}{w^{RD}},$$

which has the same value in the new equilibrium. Therefore, the unique equilibrium with inelastic labor supply is characterized by a constant ratio $\frac{\pi}{w^{RD}}$. Given that the labor supply is inelastic, L^{RD} in the above expression can be read as the labor demand for R&D:[23]

$$L^{RD,d}\left(w^{RD}\right) = \frac{\pi}{w^{RD}} - \rho\left(\mu_1 \frac{v(1)}{w^{RD}} + \mu_\omega \frac{v(\omega)}{w^{RD}}\right)$$

Now consider an initial equilibrium with $L^{RD,d}(w^{RD}) = L^d$. A change in the wage w^{RD} to $w^{RD\prime} > w^{RD}$ modifies the above expression to:

$$L^{RD,d}\left(w^{RD\prime}\right) = \frac{\pi}{w^{RD\prime}} - \rho\left(\mu_1' \frac{v'(1)}{w^{RD\prime}} + \mu_\omega' \frac{v'(\omega)}{w^{RD\prime}}\right).$$

By what shown above, it must be:

$$\frac{\mathrm{d}\pi}{\pi} = \frac{w^{RD\prime} - w^{RD}}{w^{RD}} > 0$$

for $L^{RD,d}$ to be unchanged. Thus, denoting:

$$\pi' = \pi\left(1 + \frac{w^{RD\prime} - w^{RD}}{w^{RD}}\right),$$

the above expression reads:

$$L^{RD,d}\left(w^{RD\prime}\right) = \frac{\pi'}{w^{RD\prime}} + \frac{\pi - \pi'}{w^{RD\prime}} - \rho\left(\mu_1' \frac{v'(1)}{w^{RD\prime}} + \mu_\omega' \frac{v'(\omega)}{w^{RD\prime}}\right).$$

That is:

$$L^{RD,d}\left(w^{RD\prime}\right) = L^{RD,d}\left(w^{RD}\right) + \frac{\pi - \pi'}{w^{RD\prime}} < L^{RD,d}\left(w^{RD}\right).$$

This shows that labor demand is decreasing in the wage. In general, we have:

$$L^{RD,d}\left(w^{RD\prime}\right) = L^{RD,d}\left(w^{RD}\right) + \frac{1}{w^{RD}}\left(\frac{w^{RD}}{w^{RD\prime}} - 1\right)$$

[23] Alternatively, the market clearing expression can be rewritten as the accounting identity that instantaneous profits equal the R&D wage bill plus dividends, which gives the demand for R&D labor as the expression reported below.

Consider now $w^{RD\prime} \to 0$, in this case we clearly have:

$$L^{RD,d}\left(w^{RD\prime}\right) \to \infty.$$

Conversely, with $w^{RD\prime} \to \infty$:

$$L^{RD,d}\left(w^{RD\prime}\right) \to L^{RD,d}\left(w^{RD}\right) - \frac{1}{w^{RD}} = -\rho\left(\mu_1 \frac{v(1)}{w^{RD}} + \mu_\omega \frac{v(\omega)}{w^{RD}}\right) - \frac{wL}{w^{RD}} < 0.$$

□

By Lemma 5, given an endowment of production labor and an R&D labor supply schedule, the CGE is unique.

To derive the growth rate note that, by the Cobb Douglas assumption on the final good, and given the equilibrium wage rate for production workers, $w = \frac{w_t}{Y_t}$,

$$\begin{aligned}
\log Y_t &= \int_0^1 \log y_t(i)\mathrm{d}i \\
&= \int_0^1 \log\left(\frac{Y_t}{w_t c_t(i)}\right)\mathrm{d}i \\
&= \int_0^1 \log\left(\frac{1}{w c_t(i)}\right)\mathrm{d}i.
\end{aligned}$$

It follows that:

$$\begin{aligned}
g = \log(Y_{t+\Delta t}) - \log(Y_t) &= -\int_0^1 \left(\log c_{t+\Delta}(i) - c_t(i)\right)\mathrm{d}i \\
&= \eta\left[x_{e,\omega}\mu_{e,\omega} + x_{e,1}\mu_{e,1} + \lambda x_I \mu_1\right] \\
&= \eta\left[x_{e,\omega}\left(\mu_{e,\omega} + \omega\mu_{e,1}\right) + \lambda x_I \mu_1\right].
\end{aligned}$$

Productivity g/L^{RD} follows directly from total R&D labor demand:

$$\zeta\omega x_{e,\omega}\left(\mu_{e,\omega} + \mu_{e,1}\right) + \alpha_I \frac{(x_I)^\gamma}{\gamma}\mu_1.$$

□

Proof of Proposition 3. The increase in R&D efforts by both incumbents and entrants descend directly from Lemma 4. In what follows, I derive *equilibrium* quantities, that is factoring in wage effects, but I drop stars for ease of notation.

To prove that the share of R&D labor accruing to incumbents increases, note first:

$$\frac{\partial L_I}{\partial \phi} = \alpha_I x_I^{\gamma-1} \mu_1 \frac{\partial x_I}{\partial \phi} + \frac{\alpha_I}{\gamma} x_I^{\gamma-1} \frac{\partial \left(x_I \mu_1\right)}{\partial \phi},$$

where the first term is strictly positive, since I have prover that $\frac{\partial x_I}{\partial \phi} > 0$, and the term, $\frac{\partial(x_I \mu_1)}{\partial \phi}$,denotes the derivative of aggregate incumbents' research intensity with respect to the markup, and is also strictly positive. Indeed:

$$\frac{\partial \mu_1}{\partial \phi} = \frac{\partial \left(\frac{x_{e,\omega}+\delta}{x_{e,\omega}+\delta+x_I}\right)}{\partial \phi} = \left[\frac{\frac{\partial(x_{e,\omega}+\delta)}{\partial \phi} x_I - \left(x_{e,\omega}+\delta\right) \frac{\partial x_I}{\partial \phi}}{\left(x_I + x_{e,\omega} + \delta\right)^2}\right] = \mu_1 \frac{\partial x_I}{\partial \phi} \frac{(\epsilon - 1)}{\left(x_I + x_{e,\omega} + \delta\right)}, \tag{32}$$

where I define the ratio of the elasticity of $x_{e,\omega} + \delta$ and x_I to ϕ as:

$$\epsilon \equiv \frac{\epsilon_e}{\epsilon_I} \equiv \frac{\frac{\partial(x_{e,\omega}+\delta)}{\partial \phi} / x_{e,\omega}}{\frac{\partial x_I}{\partial \phi} / x_I} \in (0,1].$$

therefore:

$$\begin{aligned}
\frac{\partial \left(\mu_1 x_I\right)}{\partial \phi} &= \mu_1 \frac{\partial x_I}{\partial \phi} \left[\frac{x_I (\epsilon - 1)}{\left(x_I + x_{e,\omega} + \delta\right)} + 1\right] \\
&= \mu_1 \frac{\partial x_I}{\partial \phi} \left[\frac{x_I \epsilon + x_{e,\omega} + \delta}{\left(x_I + x_{e,\omega} + \delta\right)}\right] > 0.
\end{aligned}$$

This proves that the aggregate incumbents' research intensity, $x_I \mu_1$, is increasing in the markup. By (32), μ_1 decreases with ϕ if and only if $\epsilon < 1$, that is, x_I is more elastic than $x_{e,\omega}$ to changes in the markup. Therefore, I now proceed to show that, when $\lambda = 0$, productivity is unambiguously decreasing in ϕ if the mass of unprotected markets, μ_1, falls with ϕ. With $\lambda = 0$, inventors' productivity reads:

$$\begin{aligned}
\frac{g}{L^{RD}} &= \eta \frac{x_{e,\omega} \left(\mu_{e,\omega} + \omega \mu_{e,1}\right)}{L_e + L_I}, \\
&= \eta \frac{x_{e,\omega} \left(\mu_{e,\omega} + \omega \mu_{e,1}\right)}{L_e \left(1 + \frac{L_I}{L_e}\right)} \\
&= \frac{\eta}{\zeta \omega} \underbrace{\frac{\mu_{e,\omega} + \omega \mu_{e,1}}{\mu_{e,\omega} + \mu_{e,1}}}_{\equiv R} \frac{1}{\left(1 + \frac{L_I}{L_e}\right)},
\end{aligned}$$

where L_e denotes entrants' R&D labor, $\zeta \omega x_{e,\omega} \left(\mu_{e,1} + \mu_{e,\omega}\right)$, and L_I denotes incumbents' inventors, $\frac{\alpha_I}{\gamma} x_I^{\gamma}$. By what I have shown above, L_I / L_e increases with ϕ, so the second term is decreasing in the markup. The statement is the verified if the first ratio, R, is also decreasing in ϕ. Dividing numerator and denominator in R by $\mu_{e,\omega}$,we have that:

$$R = \frac{1+\omega\frac{\mu_{e,1}}{\mu_{e,\omega}}}{1+\frac{\mu_{e,1}}{\mu_{e,\omega}}}.$$

Since $\omega > 1$, R increases in the ratio of entrants in unprotected versus protected markets, as intuitive. Now define this ratio writes, using the stationary distribution of entrants in (28) and (29), and after some algebra:

$$\frac{\mu_{e,1}}{\mu_{e,\omega}} = \frac{\omega\left[\frac{\mu_1}{1-\mu_1}\right]^2}{\omega\frac{\mu_1}{1-\mu_1}\left(\frac{2x_{e,\omega}+\delta}{x_{e,\omega}+\delta}\right)+1} + \frac{\delta}{\omega\left(x_{e,\omega}+\delta\right)+\omega x_{e,\omega}+x_I},$$

where the second term is always decreasing in ϕ since research intensities are increasing in ϕ. Provided that μ_1 is decreasing in ϕ, it is also straightforward to show that the first term is decreasing if μ_1 decreases.[24]

This proves that if the mass of unprotected markets, μ_1, decreases with markups, R&D productivity also falls. By (32), μ_1 decreases with ϕ if and only if $\epsilon < 1$, that is, x_I is more elastic than $x_{e,\omega}$,

$$\frac{\partial x_I}{\partial\phi}\frac{\phi}{x_I} > \frac{\partial x_{e,\omega}}{\partial\phi}\frac{\phi}{x_{e,\omega}},$$

proving the statement.[25] □

Corollary 6. *If costs are quadratic, $\gamma = 2$, there is no depreciation, $\delta = 0$, and the supply of inventors is perfectly elastic, a sufficient condition for productivity to decrease with markups is given by:*

$$\sqrt{\zeta\frac{\phi-1}{\phi}}\left(\frac{\alpha_I-\zeta\omega(\omega-1)}{\alpha_I\zeta\omega}\right) > \rho.$$

Proof. In the quadratic case, optimal incumbents' research intensity reads:

$$x_I = \frac{v(\omega)-v(1)}{\alpha_I w^{RD}}$$

[24] Let $z \equiv \frac{\mu_1}{1-\mu_1}$, $t \equiv \left(\frac{2x_{e,\omega}+\delta}{x_{e,\omega}+\delta}\right)$, and let primes denote derivatives with respect to ϕ. Then:

$$\partial\left[\frac{\omega z^2}{\omega z t+1}\right] = \frac{2\omega z z' + 2\omega^2 z^2 z' t - \omega^2 z^2 z' t - \omega z^2 t'}{(\omega z t+1)^2} = \frac{\omega^2 z^2 z' t + 2\omega z z' - \omega z^2 t'}{(\omega z t+1)^2} < 0$$

if $z' < 0$. Indeed $t' > 0$ since $x_{e,\omega}$ increases in ϕ.

[25] In particular, this condition holds if, at given wages, the elasticity of incumbents' demand for research intensity is larger than entrants', and

$$(\omega-1) \in \left[2, \frac{1}{\gamma-1}\right].$$

In this case, it is possible to show that incumbents' demand for research intensity is less wage elastic than entrants', so equilibrium wage effects do not overturn demand effects on the ratio $x_I/x_{e,\omega}$.

Therefore:

$$\partial\left[\frac{x_I}{x_{e,\omega}}\right]=\frac{\zeta\omega}{\alpha_1}\partial\left[\frac{v(\omega)}{v(1)}-1\right].$$

Therefore the elasticity of x_I to ϕ is larger than that of $x_{e,\omega}$ if and only if:

$$\operatorname{sign}\left(\frac{\partial\left(v(\omega)/v(1)\right)}{\partial m}\right)=\operatorname{sign}\left(\frac{\partial v(\omega)}{\partial m}v(1)-\frac{\partial v(1)}{\partial m}v(\omega)\right)>0. \tag{33}$$

By Lemma 4 applied to the case $\gamma=2$:

$$\begin{aligned}\begin{bmatrix}\frac{\mathrm{d}v(1)}{\mathrm{d}\pi}\\ \frac{\mathrm{d}v(\omega)}{\mathrm{d}\pi}\end{bmatrix}&=\frac{1}{\det J}\begin{bmatrix}\frac{\rho\zeta\omega+v(1)}{\zeta\omega} & \frac{v(\omega)-v(1)}{\alpha_I}\\ -\frac{v(\omega)}{\zeta\omega} & \rho+\frac{v(\omega)-v(1)}{\alpha_I}+2\frac{v(1)}{\zeta}\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\\ &=\frac{1}{\det J}\begin{bmatrix}\rho+x_{e,\omega} & x_I\\ -x_{e,\omega} & \rho+x_I+2\omega x_{e,\omega}\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}\end{aligned}$$

Thus Equation (33) has the same sign as:

$$\left(\rho+x_I+(2\omega-1)\,x_{e,\omega}\right)v(1)-\left(\rho+x_I+x_{e,\omega}\right)v(\omega).$$

With $\omega>1$, by Lemma 4 it holds:

$$\omega v(1)>v(\omega).$$

Therefore a sufficient condition for the ratio z to increase in m is:

$$\begin{aligned}\left(\rho+x_I+(2\omega-1)\,x_{e,\omega}\right)&>\omega\left(\rho+x_I+x_{e,\omega}\right)\\ (\omega-1)\,x_{e,\omega}&>(\omega-1)\left(\rho+x_I\right)\\ x_{e,\omega}-x_I&>\rho.\end{aligned}$$

Using once again, $\omega v(1)>v(\omega)$, it is possible to write

$$\begin{aligned}x_{e,\omega}-x_I&>x_{e,\omega}\left(1-\zeta\omega\frac{(\omega-1)}{\alpha_I}\right)\\ &=\frac{v(1)}{\zeta\omega}\left(1-\zeta\omega\frac{(\omega-1)}{\alpha_I}\right)\\ &=v(1)\left(\frac{\alpha_I-\zeta\omega\,(\omega-1)}{\alpha_I\zeta\omega}\right).\end{aligned}$$

Finally, by definition of the value function:

$$\rho v(1)\geq m-\frac{v(1)^2}{\zeta},$$

with equality only when it is optimal for incumbents not to invest. Solving gives:

$$v(1) > \frac{-\rho\zeta + \sqrt{(\rho\zeta)^2 + 4\zeta m}}{2} > \sqrt{\zeta m}$$

Therefore:

$$x_{e,\omega} - x_I > \sqrt{\zeta m}\left(\frac{\alpha_I - \zeta\omega(\omega-1)}{\alpha_I\zeta\omega}\right) > \rho,$$

proving that the statement gives a sufficient condition for the elasticity of x_I to be larger than $x_{e,\omega}$. By Proposition 3, it follows that when this condition is satisfied, increases in markup lower growth. □

C.2 Full Description of the Two-Sector Model and Derivations

By the above assumptions, the final good is produced according to:

$$Y = \prod Y_i^{\beta_i}. \tag{34}$$

With the final good as numeraire, the sector's demand schedule is:

$$Y_i = \beta_i \frac{Y}{P_i}. \tag{35}$$

From CD on intermediate goods we also have:

$$P_i Y_i = p_{is} y_{is}, \quad \forall s.$$

In each sector, the price is set at the competitive fringe's marginal cost wc_i, and is identical across subsectors . Thus

$$P_i = p_{is} = wc_i, \; Y_i = \beta_i \frac{Y}{wc_i}. \tag{36}$$

Equilibrium profits are given by:

$$\Pi_i = \left(c_i w - \frac{c_i w}{\phi_i}\right) Y_i = \left(\frac{\phi_i - 1}{\phi_i}\right) \beta_i Y.$$

The monopolist demands production labor:

$$\ell_{is} = \frac{c_i y_{is}}{\phi_i}, \Rightarrow L_i = \int \ell_{is} \mathrm{d}s = Y \frac{\beta_i}{\phi_i w}. \tag{37}$$

Assuming a rigid production labor supply:[26]

$$L^s(w) = L = \frac{Y}{w}\left(\sum \frac{\beta_i}{\phi_i}\right). \tag{38}$$

Which gives:

$$L_i = L\frac{\frac{\beta_i}{\phi_i}}{\sum \frac{\beta_i}{\phi_i}}, Y_i = L\frac{\frac{\beta_i}{c_i}}{\sum \frac{\beta_i}{\phi_i}}. \tag{39}$$

Which gives:

$$Y = L\prod_i \left(\frac{\frac{\beta_i}{c_i}}{\sum \frac{\beta_i}{\phi_i}}\right)^{\beta_i}. \tag{40}$$

Thus, growth is:

$$-\sum \beta_i \Delta \log c_i. \tag{41}$$

Normalized values in each sector are the same as before, with the only difference that they receive a wage w^R, and the above α_I, ζ are replaced by $\zeta w^R, \alpha_I w^R$.

C.2.1 Research Equilibrium in the two-sector model

By the above solutions, the monopolist's values read:

$$\rho V_i(1) = \max_{x_I}\left(\frac{\phi_i - 1}{\phi_i}\right)\beta_i Y - \alpha_I W^{RD}\frac{x_I^2}{2} + x_I\left(V_i(\omega) - V_i(1)\right) - x_{e,1}V_i(1),$$
$$\rho V_i(\omega_i) = \left(\frac{\phi_i - 1}{\phi_i}\right)\beta_i Y + \delta\left(V(1) - V(\omega)\right) - x_{e,\omega}V_i(\omega).$$

And normalized values, $v \equiv V/Y$:

$$\rho v_i(1) = \max_{x_I}\left(\frac{\phi_i - 1}{\phi_i}\right)\beta_i - \alpha_I w^{RD}\frac{x_I^2}{2} + x_I\left(v_i(\omega) - v_i(1)\right) - x_{e,1}v_i(1) \tag{42}$$
$$\rho v_i(\omega_i) = \left(\frac{\phi_i - 1}{\phi_i}\right)\beta_i + \delta\left(v(1) - v(\omega)\right) - x_{e,\omega}v_i(\omega), \tag{43}$$

[26] Consider a labor supply with elasticity φ. This gives:

$$\chi w^{\varphi} = \frac{Y}{w}\left(\sum \frac{\beta_i}{\phi_i c_i}\right) \Rightarrow w = \left[\frac{Y}{\chi}\left(\sum \frac{\beta_i}{\phi_i c_i}\right)\right]^{\frac{1}{1+\varphi}}$$

Equilibrium labor is then:

$$L^\star = \chi\left[\frac{Y}{\chi}\left(\sum \frac{\beta_i}{\phi_i c_i}\right)\right]^{\frac{\varphi}{1+\varphi}}, \frac{Y}{w} = Y^{\frac{\varphi}{1+\varphi}}\left[\frac{1}{\chi}\left(\sum \frac{\beta_i}{\phi_i c_i}\right)\right]^{-\frac{1}{1+\varphi}} = L^\star\left(\sum \frac{\beta_i}{\phi_i c_i}\right)^{-1}$$

Which results in the same allocations and outputs as below, with $L^\star$ in place of the fixed L.

where w^{RD} is the normalized researchers' wage.

Given a normalized wage, each sector demands:

$$x_{e,\omega,i}\left(w^{RD}\right) = \frac{v_i(1)}{w^{RD}\omega\zeta_i}, \tag{44}$$

$$x_{I,i}\left(w^{RD}\right) = \frac{(v_i(\omega_i) - v_i(1))}{w^{RD}\alpha_{I,i}}. \tag{45}$$

The stationary distribution within each sector is given by:

$$\mu_{\omega,i}\left(w^{RD}\right) = \frac{x_{I,i}\left(w^{RD}\right)}{x_{e,\omega,i}\left(w^{RD}\right) + \delta_i + x_{I,i}\left(w^{RD}\right)}, \tag{46}$$

$$\mu_{1,i}\left(w^{RD}\right) = \frac{x_{e,\omega,i}\left(w^{RD}\right) + \delta_i}{x_{e,\omega,i}\left(w^{RD}\right) + \delta_i + x_{I,i}\left(w^{RD}\right)}, \tag{47}$$

$$\mu_{e,1,i}\left(w^{RD}\right) = \frac{\omega_i\left(x_{e,\omega,i}\left(w^{RD}\right) + \delta\right)\mu_{1,i} + \delta_i\mu_{\omega,i}}{\left(x_{I,i} + \omega_i\left(x_{e,\omega,i}\left(w^{RD}\right) + \delta_i\right)\right)}, \tag{48}$$

$$\mu_{e,\omega,i}\left(w^{RD}\right) = \frac{\omega_i\mu_{1,i}x_{I,i}\left(w^{RD}\right) - \omega_i\delta_i\mu_{\omega,i}}{\left(x_{I,i} + \omega_i\left(x_{e,\omega,i}\left(w^{RD}\right) + \delta_i\right)\right)} + \mu_{\omega,i}. \tag{49}$$

Sector RD labor demand is given by:

$$L_i^{RD,d}\left(w^{RD}\right) = \mu_{e,\omega,i}\left(w^{RD}\right)\left(\zeta_i\omega_i x_{e,\omega,i}\left(w^{RD}\right)\right) + \mu_{1,e,i}\left(w^{RD}\right)\zeta_i x_{e,1,i}\left(w^{RD}\right) + \mu_{1,i}\left(w^{RD}\right)\alpha_I\frac{x_{I,i}^2\left(w^{RD}\right)}{2}.$$

With an inelastic labor supply fixed to L^{RD}, market clearing for inventors then reads:

$$L^{RD} = \sum_i\left\{\mu_{\omega,i}\left(w^{RD}\right)\left(\zeta_i\omega_i x_{e,\omega,i}\left(w^{RD}\right)\right) + \mu_{1,e,i}\left(w^{RD}\right)\zeta_i x_{e,1,i}\left(w^{RD}\right) + \mu_{1,i}\left(w^{RD}\right)\alpha_I\frac{x_{I,i}^2\left(w^{RD}\right)}{2}\right\}. \tag{50}$$