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

Andrea Manera†

September 22, 2021

PRELIMINARY AND INCOMPLETE

Abstract

Inventors are a scarce resource, whose skill sets can apply to R&D in disparate product markets. Motivated by this observation, I explore the impact of product market competition on the misallocation of inventors, and its implications for growth. First, I delineate the boundaries of "knowledge markets", employing USPTO patents data to group NAICS sectors that employ the same inventors. Second, I analyze the relation between 4-digit NAICS sectors market concentration and the share of inventors employed in R&D projects relevant to these sectors. Four findings emerge from the analysis. First, the last thirty years saw a sizable increase in concentration of inventors across both patent (CPC) classes and their 4-digit sector of application. Second, over the period 1997-2012, increases in sector-level concentration are positively correlated with the share of inventor markets captured by each sector. An IV analysis based on the increase in sector-specific regulations suggests a causal interpretation of this result. Third, sectors with increased inventor concentration have seen a decrease in the productivity of inventors, as measured by the growth in output per worker. Fourth, higher shares of relevant inventors are positively correlated with self-citations and with rising concentration of inventors at the top of the innovating firms distribution, and negatively with forward citations. A back-of-the-envelope computation suggest that the increased concentration of inventors in less competitive sectors can for up to 22.5% of the overall decrease in output per worker growth over the period 1997-2012, which corresponds to a fall of .45% in absolute value. To rationalize my findings, I propose a Schumpeterian model of creative destruction, where incumbents can conduct defensive patenting, to rationalize my findings. In the model, higher markups increase R&D effort, as well as causing its allocation to defensive projects by incumbents. When different product markets compete for the same inventors, a largest share of researchers end up employed in defensive projects in less competitive sectors, lowering growth and R&D productivity.

*I thank...

.

.

.

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

1 Introduction

There is a long-standing debate in the press and policy circles over the allocation of talent in the economy, and whether skilled work is employed where it is most beneficial to economic growth and society overall. Central in the current policy debate is the concern that dominant big-tech firms are inefficiently appropriating a large share of highly-educated, highly-skilled workers, subtracting this crucial resource from more competitive and dynamic sectors. A stark example of this phenomenon is the high concentration of AI expert workforce. TalentSeer (2020) estimates that up to 20% of total AI experts are employed by just five companies: Google, Microsoft, Apple, Amazon and IBM. At the same time, smaller firms in other economic sectors appear unable to attract this talent, raising concerns over the potential misallocation of innovators equipped with the latest technological skills.

This paper studies the broader effect of increasing concentration on the allocation of inventors across sectors, and shows that such concerns are justified. Using USPTO patent data and concentration measures from the Economic Census over the period 1997-2017, I demonstrate that sectors where concentration has increased have indeed attracted an increasing share of inventors. This relation has a causal interpretation, according to an IV specification where I use the increase in the number of regulations from Mercatus RegData as an instrument for increased concentration. I establish three additional facts that suggest that inventors are misallocated. First, researchers accrued mostly to incumbent firms in concentrated sectors. Second, the productivity of R&D in these sectors has fallen. Third, the patents deposited in these sectors have seen more self-citations and less forward citations, suggesting that additional inventors have been allocated to incremental project that did not pave the way for further developments. Quantitatively, my findings imply that inventor misallocation led to a fall of up to .45pp in output-per-worker growth in the sectors I study (20% of the overall observed reduction).

Based on these four observations, the main hypothesis of this paper is that differences in concentration across product markets that employ the same types of inventors generate a inefficiency, which pushes economic growth below its potential. Indeed, inventors are misallocated towards less competitive markets where defensive projects—which hamper entry and Schumpeterian growth—are more prevalent, and away from more competitive sectors where growth is more sustained.

I offer three main contributions. First, I construct a dataset of "inventor markets", defined as collections of product markets that share similar inventors. These inventor markets originate from the network generated by transitions of individual inventors across product categories, identified using USPTO patent data, where patents are classified according to their main NAICS sector of applications.

Second, I provide empirical evidence in favor of the misallocation channel described above, using USPTO patent data, Economic Census concentration measures, and Mercatus regulation data over the 1997-2017 horizon. I show that, within inventor markets, sectors where competition has fallen increased their share of researchers. In addition, inventors have accrued to incumbent firms in less competitive sectors, where show that growth in output per worker has fallen, self-citations have increased, and forward citations decreased. These results suggests that researchers have been increasingly allocated to defensive projects. Quantitatively, my findings imply that increased market concentration over the last 20 years can explain up to 20% of the overall fall in inventor productivity, which many studies identified as a culprit for stagnant overall productivity growth.

Third, in order to interpret these results, I build a Schumpeterian model where, in addition to productive R&D conducted by new entrants, incumbents can engage in defensive projects to reduce entry and strengthen their dominant position. A two-sector general equilibrium model shows that unbalanced changes in concentration across sectors generate a fall in inventor productivity and growth

CONJECTURE: Under the lenses of my model, a planner wishing to maximize long-run growth and taking the market structure as given, should provide a sector-specific R&D subsidy to entrants in order to reduce the impact of defensive patent walls and to conform the sectoral R&D allocation to maximize aggregate GDP growth.

2 Data description

2.1 Data Sources

- *Patents.* PatentsView USPTO, PATSTAT patent classification into NAICS by Goldschlag et al. (2016), crosswalk between the two sectors built by Gianluca Tarasconi (2019).[1]
- *Concentration and Sales.* US Census extended data from Keil (2017).[2]
- *Price Indices.* To deflate sales, I use NAICS-specific price indices from the BLS.
- *Market Regulations.* Mercatus RegData 4.0.[3]
- *Productivity.* I use output per worker from the Economic Census.

2.2 Constructed Data

- *Knowledge Markets.* I build knowledge markets using observed flows of inventors across projects classified through text analysis into different NAICS 4-digit codes by Goldschlag et al. (2016). I maximize the modularity of the resulting network using the Louvain method to identify communities of NAICS that are connected by inventor flows over the period 1976-2015. I obtain about 10 non-singleton markets.

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

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

[3] https://www.quantgov.org/bulk-download.

- *Effective Inventors.* As discussed above, I compute inventor productivity as the inventor fixed effect α_i from the fully-saturated regression:

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

  at the level of CPC (Cooperative Patent Classification) class, c, assignee/company, f, inventor i and year t. I use the broadest CPC class to maximize the number of fixed effects that I can identify. The term α_{cft} is a CPC class by firm by time fixed effect.

- *Extended Regulation Series.* Mercatus RegData provides a count of restrictions imposed on a number of NAICS 4d-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 for unspecified reasons.
  Therefore, I process the description of NAICS 4d 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. In particular, 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 (XXX provide examples in footnote XXX).

- *Share of self-citations and excess citation measures.* For each patent classified by **?**, i.e. with non-missing NAICS classification, I count the set of cited patents that belong to the citing patent's assignee. In the case of cited patents with multiple assignees, I consider half a count if the assignee is among them. The share of self-citation is given by this count divided by total citations. 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 various measures differ 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. In employ in turn the share of NAICS patents in the year, 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 constructed for each citation by taking the share of patents by the assignee in the CPC (sub)group and year corresponding to the cited patents.[4]
  Finally, I aggregate all measures across assignees in the same NAICS 4-digit code using the number of patents in the relevant NAICS code by each assignee in each year.

- *Inventor productivity.* As a measure of inventor productivity, I use the average growth in output

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

per worker divided by inventor's fixed effect. I choose this measure since it is the productivity measure available for most sectors.

- *Forward citations and patent generality.* See Hall et al. (2000) and Acemoglu et al. (Forthcoming).

2.2.1 Aggregation at Five-Year Census Frequency

Data from the Economic Census are available at five-year frequency 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). In the IV regression I use product restrictions as an instrument for concentration, which motivates me to average restrictions in the five-year window *ending* in the census year (e.g., 1993-1997 for 1997). Since **?**'s matching only covers the period up to 2016, I run all specifications in long-differences over the time frame 1997-2012.

3 Results

I present four main findings that apply to the period 1997-2012:

1. Effective inventors have become more concentrated in specific technology classes and economic sectors;

2. Sectors with increased concentration have attracted a growing share of relevant inventor types. The IV analysis suggest that the rise in inventor shares is the result, and not the cause, of increased concentration;

3. Growth in the share of relevant inventors is negatively correlated with inventor productivity, as measured by average growth in output per workers divided by effective inventors employed;

4. Growth in the share of relevant inventors is positively correlated with the share of self-citations and excess self-citations, as well as concentration of inventors at the top within sectors.

The first finding emerges from the computation of Gini coefficients of effective inventors across technologies and sectors. Findings 2-4 come from long-difference regressions over the period 1997-2012. Regressions are weighted by sector sales in 2012 for findings 2-3, which rely on Census sector-level measures, with robust standard errors. I present both unweighted and unweighted results for finding 4, since these variables do not represent aggregates according to sales, and rely on patent data only.

For findings 2-4, I run long-difference regressions at the product market level for the period 1997-2012:

$$\Delta \text{Outcome}_p = \beta \Delta \text{Indep.Var.}_p + \gamma' \Delta \text{Controls}_p + \varepsilon_p, \tag{2}$$

where Δ denotes the long-difference operator. Throughout the analysis, I use the change in log-real sector sales or the change in log- real sector sales per company as controls, in order to capture effects on the outcome stemming from increases in the real size of sectors or firms. I also run the within-knowledge-market specification:

$$\Delta \text{Outcome}_p = f_k \mathbf{1}\{p \in k\} + \beta \Delta \text{Indep.Var.}_p + \gamma' \Delta \text{Controls}_p + \varepsilon_p, \tag{3}$$

which includes fixed effects to denote the membership of product markets, p, to the relevant knowledge market k.

3.1 Increased Effective Inventors Concentration across Patent Classes and NAICS sectors

First, I compute Gini coefficients of effective inventors, α_i (shifted to be nonnegative) across patent classes. This is reported in Figure 1. Second, I use the subset of patents classified by Zolas (which ends in 2016) that I can match to PatentsView. This is shown in Figure 2. In both cases the coefficient increased by about 10% from 1978.

Figure 1: Effective inventor Gini coefficients at CPC-4 level

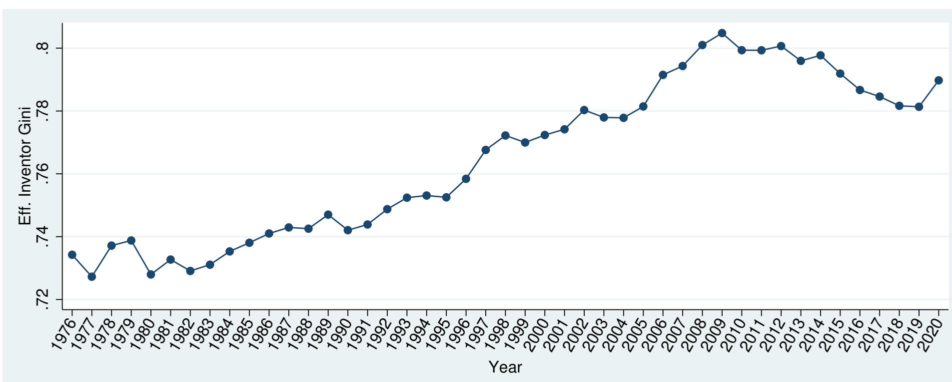
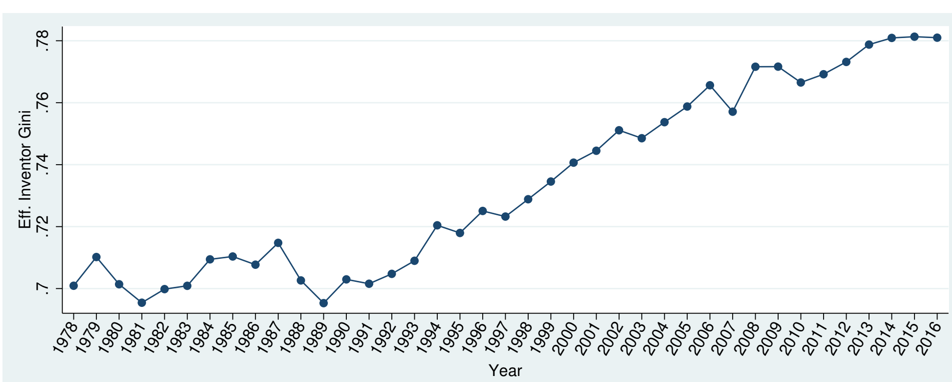


Figure 2: Effective inventor Gini coefficients at NAICS 4-digit level



3.2 Markets with Growing Concentration Increased Their Inventor Share

I compute inventor markets as described above, on the set of all patents that are assigned a NAICS code by Zolas et al's (2016) procedure. The share of effective knowledge-market inventors employed in each product market in year t is then given by the sum of inventor fixed effects from regression 1 for inventors patenting in NAICS sector, p, over their knowledge-market total:

$$\text{Share}^k_{p,t} \equiv \frac{\sum_{i\in p(t)} \alpha_i}{\sum_{i\in k(t)} \alpha_i}.$$

Here, the notation $i \in p(t)$ indicates that inventor i patents in product market p at time t, while $i \in k(t)$ denotes that inventor i patents in knowledge market k at time t. Two notes are in order. First, by construction, each product market, p, belongs to only one knowledge market, k. Second, inventors can patent in several product markets in each year. In this case, I attribute the inventor fixed-effect α_i to all product and knowledge markets to which the inventor contributes. The total effective inventor share is defined analogously as

$$\text{Share}_{p,t} \equiv \frac{\sum_{i\in p(t)} \alpha_i}{\sum_i \alpha_i}.$$

In this section, I present three sets of results for each specification, which differs 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 which trims the most extreme observations for the outcome and the independent variable separately. I trim 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.[5] "Mahalanobis 5%" denotes the sample where I trim 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 outcome and independent variable, the sample varies in each regression.[6]

Table 1 presents the results of regression 5 where the outcome variable is the change in knowledge-market inventor share, and the independent variable is the change in the lower bound of the Herfindal-Hirschman Index, or the index as reported by the Economic Census. The Economic Census reports HHI indices only for a subset of NAICS 4-digit sectors, but includes concentration ratios for a much wider set. These concentration ratios are used by Keil (2017) to construct a lower bound on the HHI, which I employ as my main measure of concentration due to its wider availability. The results in Table 1 highlight a strongly significant positive correlation between the change in HHI and the change in the share of effective inventors accruing to each NAICS sector. Note that this regression is only partially

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

[6] I like this but it does not always produce the results one would expect. I am thinking of implementing a version where I trim the residualized variables. Or experiment with difference distances.

driven by the contemporaneous correlation between the two variables. As discussed above, the share of effective inventors is average 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 which refer to shares or growth rates are reported in percentage points for ease of interpretation. Therefore, for example the coefficient in Column (1) of Table 1 should be interpreted as saying that an increase in one unit of the HHI index leads to an increase in the share of the relevant knowledge market of 27.25 pp. Second, HHI indices are instead constructed to range between 0 and 1. In particular, the HHI lower bound has sales-weighted an average of about .03, and a standard deviation of .032 in 2012. According to Table 1, a standard deviation increase in this measure is associated to a 0.87 pp increase in the share of inventors accruing to the relevant NAICS sector. This 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. Clearly, 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, so I use the HHI lower bound in most analyses below.

While suggestive, the correlation presented above is far from ideal, as it 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 in a specific sector. Second, it estimates the correlation both across and within knowledge markets. In Table 2, I address these two limitations by restricting the analysis within knowledge markets, and controlling for two measures of size. In the upper panel of Table 2, I use the change in the logarithm of real sales as a measure of the size of each sector, while in the lower panel I present the results when average sales per firm are included as a control. The inclusion of sales per firm is motivated by the fact that there might be significant barriers to entry to R&D, 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 the upper panel. The results in Table 2 confirm the positive relation between the change in inventor shares and concentration, and are largely unchanged relative to the estimates in Table 1, 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 implied by their expansion in overall sales as well as average firm size.

Figure 3 depicts graphically the residualized observations underlying the estimated coefficients in Columns (2) and (6) of Table 2, 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, with the marker size proportional to the regression weight. Although the sample displays some observations that appear extreme, the bulk of observations—and especially weighted observations—falls on the regression lines, mitigating the 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, both manually and defining extreme observations based on 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 variable. The bottom panel of Figure 3 reports the binned scatter plot corresponding to the sample where the 5% extreme observations according to the Mahalanobis distance have been removed, and confirms that the positive relation between concentration and inventor shares is not driven by a few extreme observations. In particular, the corresponding regression results in Table 2(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. TODO: Add Appendix.

Table 1: Regressions of Change in 4-digit Knowledge Market Share over Change in HHI Measures, Long-Differences, 1997-2012

	Ch. 4d K.M. Eff. Inv. Share (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Ch. HHI lower bound	27.293*		27.183*		27.326*	
	(11.569)		(11.941)		(11.620)	
Ch. HHI		22.399***		22.399***		22.350***
		(6.345)		(6.345)		(6.343)
4D Knowledge Market						
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 ($+\ p < 0.1,^{*}\ p < 0.05,^{**}\ p < .01,^{***}\ p < .001$); 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 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 2: 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

	Ch. 4d K.M. Eff. Inv. Share (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Ch. HHI lower bound	26.093*	22.509*	25.904*	22.716*	26.111*	22.554*
	(10.696)	(10.848)	(11.124)	(10.948)	(10.725)	(11.019)
Ch. Log Real 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)
4D 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

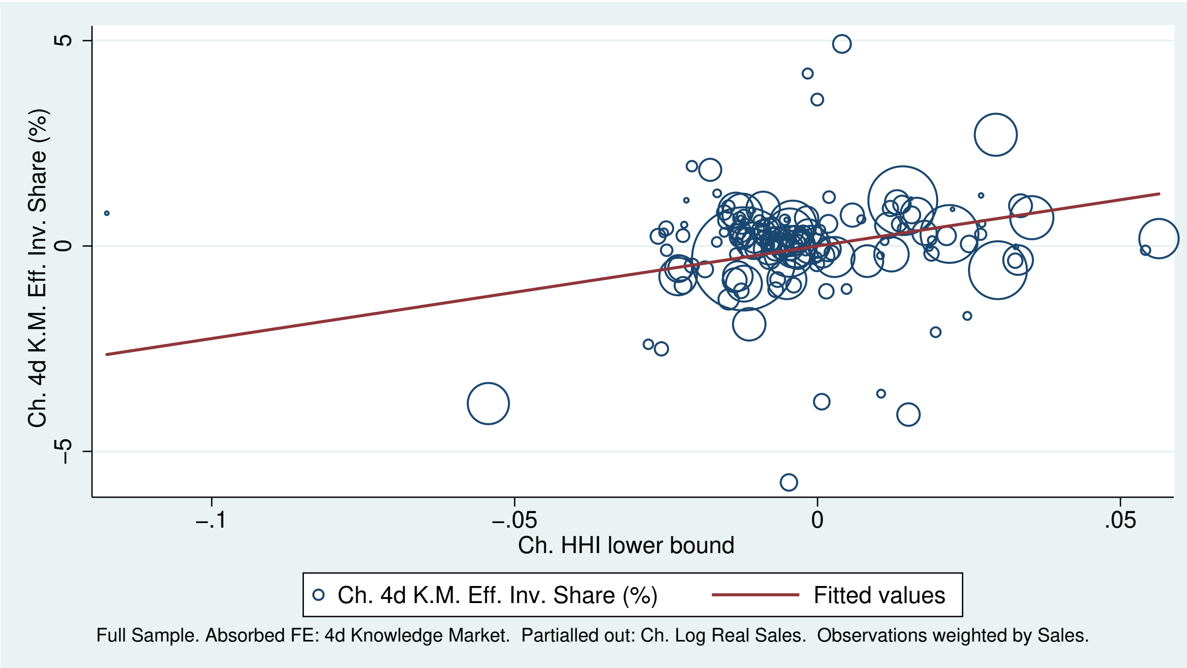
(b) Controlling for Change in Log Real Sales per Company

	Ch. 4d K.M. Eff. Inv. Share (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Ch. HHI lower bound	35.230**	20.783+	35.230**	20.783+	35.154**	22.854*
	(12.759)	(10.615)	(12.759)	(10.615)	(12.647)	(11.197)
Ch. Log Real Sales per company	0.175	-0.040	0.175	-0.040	0.300	-0.055
	(0.382)	(0.253)	(0.382)	(0.253)	(0.460)	(0.346)
4D 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

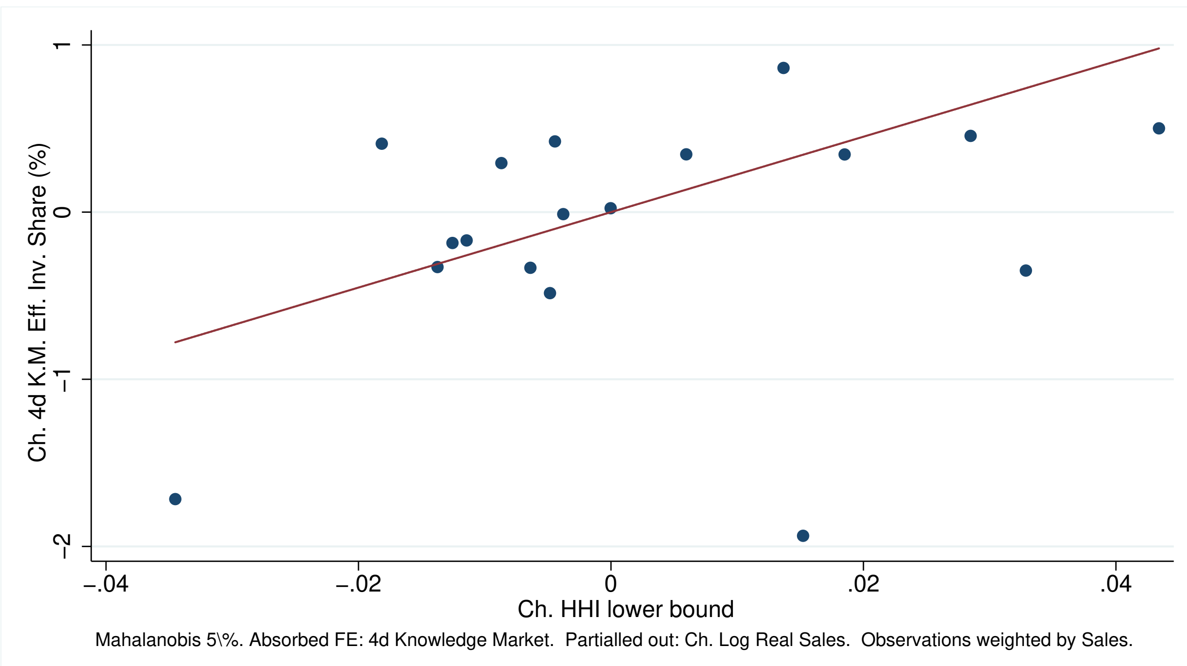
Note: Regressions weighted by sales in 2012; Robust standard errors in parentheses; Symbols denote significance levels $\left(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001\right)$; Checkmarks indicate the inclusion of fixed effects. This Tables presents the results of specifications (2) and (3), 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 3: Residualized Scatter Plots Corresponding to Selected Columns in Table 2, 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 2. 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 2.

3.2.1 IV Results

In this subsection, I present IV results that suggest that the relation between concentration and inventor shares is causal. Indeed, more concentration might just be the result of an increase in technological entry barriers, established by incumbents through an increase in their 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 the the average share of inventors following the Economic Census years to which the HHI refers as my outcome variable. However, reverse causality could still be present if the autocorrelation of inventor shares is sufficiently high. This motivates me to produce 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 the affected product markets, thus leading to higher concentration. As discussed below, this proves to be the case empirically, making a case for the validity of 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, which acts independently of product market concentration. A possibility in this sense is the increase in the number of inventors required to fulfill product market restrictions, if such regulations specifically affect technologies currently in use in the industry. However, this effect should be both large and persistent to be captured by my measure of inventor shares. Further, while RegData certainly include product restrictions, there are also a number of regulatory burdens that are not related to technological components, like reporting obligations and other legal burdens. 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 also increase. These considerations lead me to believe that the exclusion restriction is not highly likely to be violated.

The results of the 2SLS estimation are presented in the upper panel of Table 3. The specification is the same as in Column (2) of 2, 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 constructed and therefore imprecise. In particular, I had to impute regulations for a large part of the sample using the cosine-similarity between product market restrictions.[7] However, instruments are not irrelevant. The results in the lower panel of Table 3 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 equally highly significant. Accordingly, the SW underidentification test rejects the null hypothesis at a 5%

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

confidence level. Given the weakness of the instruments, I also report the Anderson-Rubin p-value, which confirms that the coefficient is 5% significant.

3.2.2 Results on Overall Inventor Shares

I also inspect 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. Results are displayed in Table 4. This shows that the relation highlighted above is not apparent when looking at the overall share of scientists, while it emerges clearly only when looking at economic sectors that actually compete for the same inventors, that is, those that belong to the same knowledge market.

Table 3: IV Regressions of Change in 4-digit Knowledge Market Share over Change in HHI Lower Bound, 2SLS Long-Difference, 1997-2012

(a) 2SLS Results

	Ch. 4d K.M. Eff. Inv. Share (%)	
	(1)	(2)
Ch. HHI lower bound	30.560+	30.096+
	(15.904)	(15.819)
Ch. Log Real Sales	0.544*	0.525*
	(0.244)	(0.247)
4D Knowledge Market FE	✓	✓
Sample	Full Sample	Mahalanobis 5%
Weight	Sales	Sales
Observations	157	150
First-Stage F	4.587229	4.753009
Anderson-Rubin p-value	.0281448	.0321185

(b) First Stage and Reduced Form

	Ch. 4d K.M. Eff. Inv. Share (%)	Ch. HHI lower bound
	(1)	(2)
Ch. Log Restricitions (NAICS 4d)	0.478*	0.016*
	(0.220)	(0.007)
Ch. Log Real Sales	0.539+	-0.000
	(0.274)	(0.005)
4D 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 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 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.

Table 4: 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)
Ch. HHI lower bound	0.297	1.692	1.328*	1.532*	0.271	1.889
	(2.007)	(1.956)	(0.649)	(0.696)	(2.038)	(2.023)
Ch. Log Real 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)
4D 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 2 for further details.

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

Table 5 presents the results of running regression (3) 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 EC year, and I proceed analogously to 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 effective inventors. Both the outcome and the dependent variable are measured in percentage points. Table 5 reveals a negative and significant correlation between the increase in the effective inventors' change and inventor productivity. These findings are robust to considering only sectors with positive growth in output per worker over the period 2012-2016. The magnitude of estimated coefficients can be grasped considering the scale of the variables and their changes over the sample period. In particular, the median change in the share of effective inventors over the period was .014%, while the measure of effective inventors has a median of 2018.[8] Using the coefficient in Column (5) to predict the median annual change in labor productivity growth implied by rising inventor concentration amounts to a fall of .15% $(-.005 \times .014\% \times 2018)$. This number increases to . 28% when using the statistics relative to sectors with positive growth in labor productivity only, 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 2(a), and given a median change in the HHI of 0.002 yields an increase in the share of effective inventors in concentrating sectors of 0.045%, which implies a fall in average labor productivity implied by misallocation of .45%. While these numbers might appear sizable considering the entirety of the economy, it is worth noting that the sample I have data for includes mainly manufacturing and retail sectors, which experienced a sizable reduction of about 2% in average

[8] 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.

annual productivity growth from 1997-2012, driven by a steep decline in output per worker growth in manufacturing. Therefore, the mechanism I propose would explain from 7.5% to 22.5% of the observed decrease in output per worker growth.

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

	Ch. Avg. Output/Worker Growth/Inventor (%)			
	(1)	(2)	(3)	(4)
Ch. 4d K.M. Eff. Inv. Share (%)	-0.007**	-0.005*	-0.007**	-0.005*
	(0.002)	(0.002)	(0.002)	(0.002)
Ch. Log Real Sales		-0.051*		-0.054*
		(0.021)		(0.021)
4D Knowledge Market FE	✓	✓	✓	✓
Sample	Full Sample	Full Sample	Mahalanobis 5%	Mahalanobis 5%
Weight	Sales	Sales	Sales	Sales
Observations	101	101	96	93

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 2 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.

3.4 Markets with Growing Inventor Shares Saw an Increase in Self-Citations and in Inventor Concentration

While the findings presented so far establish a connection between the increase in inventor concentration and the fall in inventor productivity, they shed little light on the mechanisms underlying these developments. Table 6 offers a step in this direction, showing that the increase in inventor concentration is positively correlated with the increase in self-citations within the affected sectors. This relation is apparent when restricting attention to the middle range of changes in inventor shares, reported in the Panel (b). Figure 4 clarifies why this restriction is necessary, as the full sample present some extreme observations with little change in self-citations that drive the estimated coefficient towards 0. When restricting to the middle range, which effectively consists in dropping less than 10% extreme observations, a strong positive correlation emerges between change in inventor shares and excess citations. Importantly, the measures of excess self-citations that I construct account for the contribution of active firms to technological advances in the relevant CPC classifications, thus excluding a mechanical increase in self-citations that would result solely from a reduction in R&D activity by competing firms. That is, if a firm contributes 100% of the relevant patents to a filed and cites itself only, its excess self-citations are 0, as explained in Section 2.2.

The rise in self-citations suggests an increasing role of defensive patenting by large incumbents in concentrating sectors. This explanation would also speak to the related finding that sectors that increased their inventor share also saw a *within-sector* increase in inventor concentration. Table 7 shows that the share of effective inventors accruing to top inventor-hiring firms has increased in those sectors that attracted more inventors over the period considered. This finding is consistent across a variety of measures, and Columns (4) and (5) suggests that it is driven by a faster increase of inventors at the top of the distribution more than a transfer from bottom to top firms within the sector.

Table 8 presents regressions of measures of forward citations corrected for truncation, as well as patent generality (both constructed as in Hall et al., 2000) over changes in the inventor market share. The results highlight that sectors increasing their share of inventors have experience a significant fall in forward citations per patent. As for the results in 6, the effects are more pronounced in the middle range of inventor changes. In the restricted sample, the negative correlation between inventor shares and generality is also highly significant.

Table 6: Regressions of Change in Excess Self-Citations over 4-digit Knowledge Market Share, Long-Differences, 1997-2012

(a) Full sample

	Ch. excess self-c. over CPC group		Ch. excess self-c. over CPC subgroup	
	(1)	(2)	(3)	(4)
Ch. 4d K.M. Eff. Inv. Share (%)	0.920	-0.444	0.958+	-0.228
	(0.711)	(1.083)	(0.512)	(0.801)
Ch. Log Real Sales	-1.841	-1.954	-1.456	-1.674
	(1.925)	(1.988)	(1.326)	(1.279)
4D Knowledge Market FE		✓		✓
Sample	Full Sample	Full Sample	Full Sample	Full Sample
Weight				
Observations	157	153	157	153

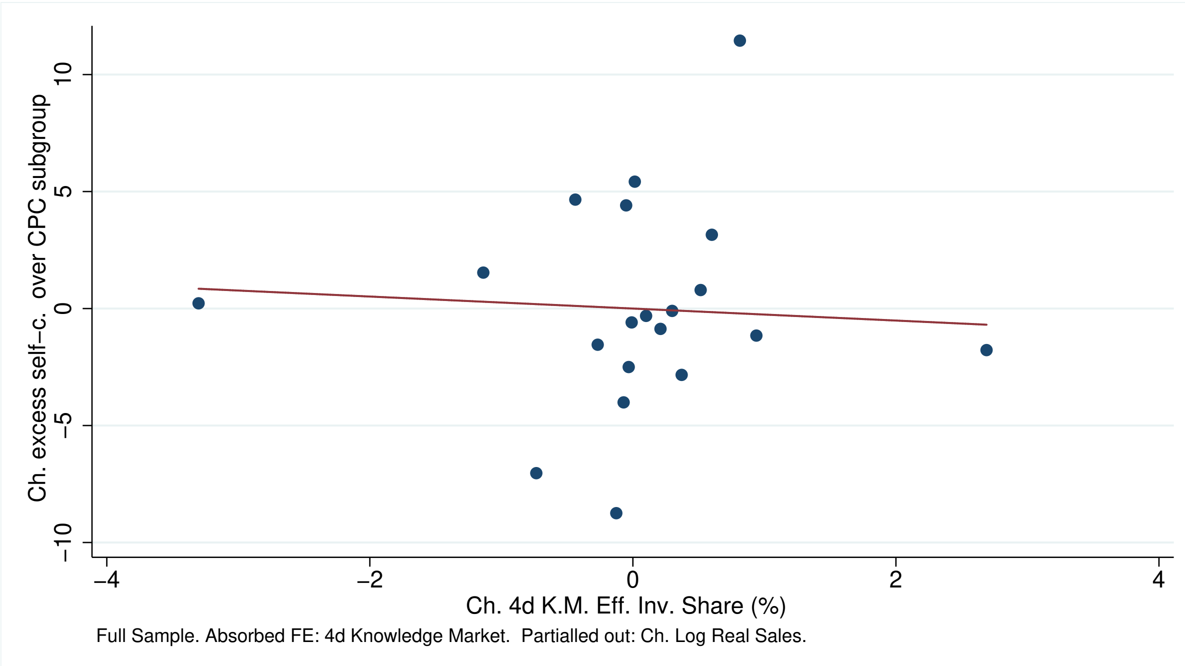
(b) Full sample, restricting to the middle range of the change in inventor shares (-2% to $+2\%$)

	Ch. excess self-c. over CPC group		Ch. excess self-c. over CPC subgroup	
	(1)	(2)	(3)	(4)
Ch. 4d K.M. Eff. Inv. Share (%)	5.540**	5.244*	4.561***	4.110*
	(1.783)	(2.469)	(1.211)	(1.600)
Ch. Log Real Sales	-2.217	-2.099	-1.780	-1.780
	(1.879)	(1.976)	(1.287)	(1.265)
4D Knowledge Market FE		✓		✓
Sample	Full Sample	Full Sample	Full Sample	Full Sample
Weight				
Observations	145	144	145	144

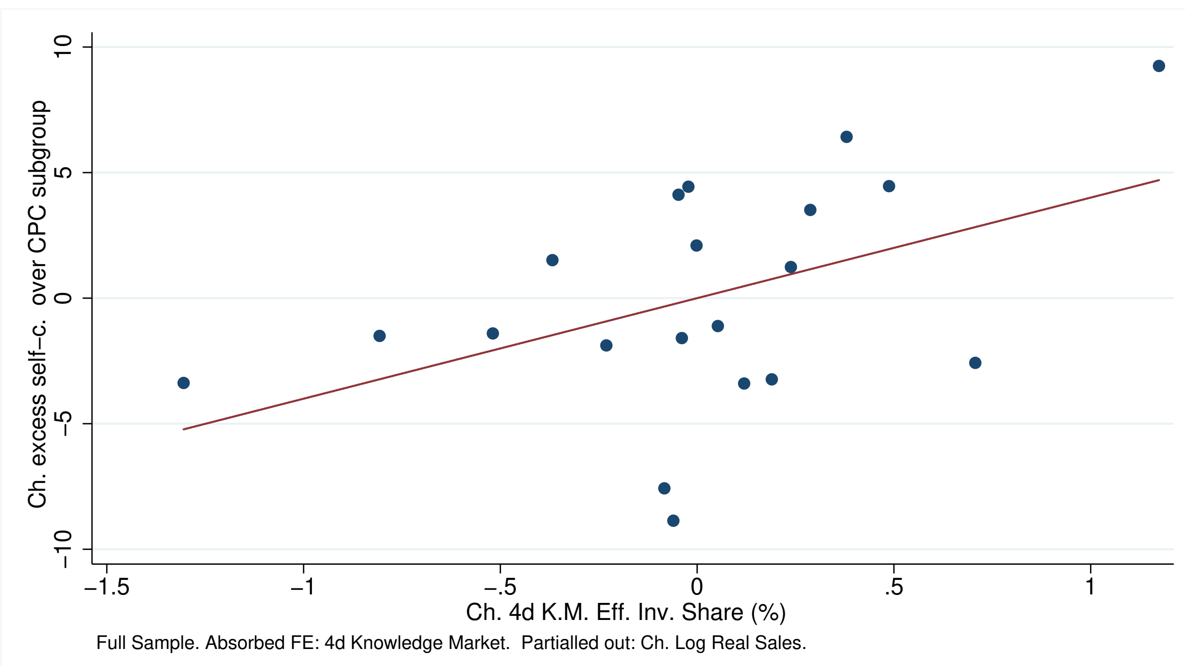
Note: Unweighted regressions; Robust standard errors in parentheses; Symbols denote significance levels $\left(+\ p < 0.1,^{*}\ p < 0.05,^{**}\ p < .01,^{***}\ p < .001\right)$; Checkmarks indicate the inclusion of fixed effects. This Tables presents the results of specifications (2) and (3), when the outcome is the change in excess self-citations in sector p over the change in the share of inventors employed in sector p. Upper panel: full sample; Bottom panel: excluding sectors with absolute increase in the inventor share above 2%.

Figure 4: Residualized Scatter Plots Corresponding to Selected Columns in Table 6

(a) Binned Scatter Plot, Full Sample



(b) Binned Scatter Plot, Middle Range of the Change in Inventor Shares (−2% to +2%)



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 2. 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 Columns (4) of the upper and bottom panel of Table 6, respectively.

Table 7: Regressions of Change in Inventor Distribution Measures over Change in 4-digit Knowledge Market Share, Long-Difference, 1997-2012

	Ch. Inv. 90/50 Quantile Ratio	Ch. Inv. Top-10/Bottom-50 Share Ratio	Ch. Inv. Top-50/Bottom-50 Share Ratio	Ch. Inv. Top 10% Share	Ch. Inv. Bottom 50% Share
	(1)	(2)	(3)	(4)	(5)
Ch. 4d K.M. Eff. Inv. Share (%)	0.211+	0.243*	0.314+	0.018**	-0.008*
	(0.107)	(0.097)	(0.184)	(0.006)	(0.004)
Ch. Log Real Sales	-0.100	0.328	0.147	0.026	0.005
	(0.122)	(0.294)	(0.316)	(0.020)	(0.007)
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 2 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 8: Regressions of Changes in Forward Citation over 4-digit Knowledge Market Share, Long-Differences, 1997-2012

(a) Full sample

	Ch. in log citations per patent (CPC2 based)	Ch. in log citations per patent (Total)	Ch. in patent generality
	(1)	(2)	(3)
Ch. 4d K.M. Eff. Inv. Share (%)	-0.197***	-0.227***	-0.004
	(0.044)	(0.051)	(0.004)
Ch. Log Real Sales	-0.234*	-0.258+	0.008
	(0.112)	(0.148)	(0.013)
4D 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 $+2\%$)

	Ch. in log citations per patent (CPC2 based)	Ch. in log citations per patent (Total)	Ch. in patent generality
	(1)	(2)	(3)
Ch. 4d K.M. Eff. Inv. Share (%)	-0.545***	-0.618***	-0.025*
	(0.113)	(0.137)	(0.012)
Ch. Log Real Sales	-0.232*	-0.255+	0.008
	(0.109)	(0.146)	(0.012)
4D 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 Tables presents the results of specification (3), 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%.

4 Model

This section presents a Schumpeterian model 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. In what follows, I first present a single-sector model to clarify the mechanism at play within each sector in the economy, and show that increases in the markup, capturing concentration in reduced form, lead to a larger amount of resources devoted to defensive innovation. Then, I move to consider a two-sector model, where each sector is identical to the single-sector model. 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. In order to illustrate the properties of the demand for inventors, I first analyze a single-sector partial-equilibrium version of the model, where labor supply of inventors and production is perfectly elastic at a fixed wage. I then consider a two-sector model where wages are determined in equilibrium, and the supply of inventors is assumed to be fixed in each instant. Proofs are reported in Appendix D.

In the empirical analysis above I used the HHI as a measure of concentration and market power. Appendix E.2 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 that adopt in what follows.

4.1 Environment: a Single-sector Model

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

$$Y_t = \int_0^1 y_t(i)\mathrm{d}i$$

The consumption good in the economy is taken as the numeraire. The market for each product $y_t(i)$ consists of a monopolist and a fringe of competitors. In what follows, I focus on a single market, dropping the argument i. Competitors can produce using a linear technology with labor requirement c_t, while the incumbent has labor requirement $\frac{c_t}{\phi}$, with $\phi > 1$. In this section, the labor supply is assumed perfectly elastic at $w = 1$. Given these assumption, the monopolist sets the limit price $p_t = c_t$ realizing profits:

$$\Pi_t = \left(c_t - \frac{c_t}{\phi}\right) y_t = \left(\frac{\phi - 1}{\phi}\right) c_t y_t,$$

By the Cobb-Douglas assumption on the final good, the demand for each product is:

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

Giving the normalized profit:

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

4.1.1 Innovation

Suppose that competitors can realize an innovation reducing unit costs to $\frac{c_t}{(1+\eta)\phi}$ investing in inventors at unit cost ζ. When one of the competitors succeed in replacing the incumbent monopolist, the other competitors can immediately realize a cost reduction to $\frac{c_t}{1+\eta}$, so that the gap between monopolist and followers is fixed at ϕ regardless of followers' innovations.

Incumbents can in turn invest in a defensive technology that lowers the productivity of followers' inventors, by a factor $\omega > 1$. I assume that this investment is completely unproductive, capturing a defensive patent that is unused by the incumbent. This defensive investment can be done only once, and is lost at rate δ (the degree of patent protection). The incumbents' problem in normalized units as a function of the state capturing entrant costs is then:

$$\begin{aligned}\rho v(1) &= \max_{x_I}\left(\frac{\phi_i - 1}{\phi_i}\right) - \alpha_I \frac{x_I^2}{2} + x_I\left(v(\omega) - v(1)\right) - x_{e,1}\left(v(1)\right) \\ \rho v(\omega) &= \left(\frac{\phi_i - 1}{\phi_i}\right) + \delta\left(v(1) - v(\omega)\right) - x_{e,\omega}\left(v(\omega)\right) \qquad (4)\end{aligned}$$

When successful, entrants obtains the unprotected monopoly position, that is, they destroy any existing patent barrier and obtain $v(1)$. I assume that the research costs of each atomistic entrant are linear in the individual research effort, as well in the innovation intensity of other entrants, capturing crowding externalities:

$$C\left(x_{e,\omega,i}\right) = \zeta\omega x_{e,\omega,i} x_{e,\omega}.$$

The maximization problem of an entrant in each state ω is then:

$$\max_{x_{e,\omega,i}} x_{e,\omega,i} v(1) - \zeta\omega x_{e,\omega,i} x_{e,\omega}.$$

The total research effort by entrants in the sector is then:

$$x_{e,\omega} = \frac{v(1)}{\omega\zeta},$$

With quadratic costs, the optimal investment for the incumbent is:

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

Lemma 4.1. *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, the normalized values,* $v(1), v(\omega)$, *are increasing in the markup factor* $m_i = \frac{\phi_i - 1}{\phi_i}$, *and*

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

4.1.2 Stationary distribution

The law of motion of the distribution across states satisfies:

$$\dot{\mu}_1 = -\left(x_I + x_{e,1}\right)\mu_1 + \delta\mu_\omega + x_{e,\omega}\mu_{e,\omega} + x_{e,1}\mu_{e,1}, \tag{5}$$

$$\dot{\mu}_\omega = -\left(x_{e,\omega} + \delta\right)\mu_\omega + x_I\mu_1, \tag{6}$$

$$\dot{\mu}_{e,1} = -\left(x_{e,1} + x_I\right)\mu_{e,1} + x_{e,1}\mu_1 + \delta\mu_{e,\omega}, \tag{7}$$

$$\dot{\mu}_{e,\omega} = -\left(x_{e,\omega} + \delta\right)\mu_{e,\omega} + x_{e,\omega}\mu_\omega + x_I\mu_{e,1}, \tag{8}$$

where $\mu_{e,1}, \mu_{e,\omega}$ are the mass of entrants targeting products without a patent barrier, and with a patent barrier, respectively. Outflows from incumbent state 1 are given by the sum of successful incumbent defensive research, and elimination of incumbents by entrants. Inflows are given by expiration of patent protection moving incumbents in state ω to state 1, and successful entry. The incumbent state ω loses mass from successful entrants and patent depreciation, and gains mass from successful defensive RD. The last two equations close the system by acknowledging that entrants lose mass from successful entry, and gain mass from displaced incumbents. The stationary distribution satisfies:

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

4.1.3 Total R&D

Total R&D intensity is given by:

$$\mu_{e,\omega}x_{e,\omega} + \mu_{e,1}x_{e,1} + \mu_1 x_I$$

By the optimal solution for entrants:

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

so total R&D effort is:

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

4.1.4 Growth and R&D productivity

Note that productivity growth is accomplished solely through creative destruction by entrants. By Cobb-Douglas on the final good:

$$\log Y_t = \int_0^1 \log y_t \mathrm{d}i$$
$$g = \log(Y_{t+\Delta t}) - \log(Y_t) = \eta \left[x_{e,\omega}\mu_{e,\omega} + x_{e,1}\mu_{e,1} \right].$$

with inventors' productivity given by:

$$\frac{g}{L^{RD}} = \eta \frac{x_{e,\omega}\left(\mu_{e,\omega} + \omega\mu_{e,1}\right)}{\zeta\omega x_{e,\omega}\left(\mu_{e,\omega} + \mu_{e,1}\right) + \alpha_I \frac{x_I^2}{2}\mu_1}.$$

4.1.5 Comparative statics in partial equilibrium

The following proposition summarizes the comparative statics in partial equilibrium. [9]

Proposition 4.2. *Consider the above model with a perfectly elastic production and R&D labor supply. Assume $\delta = 0$, and*

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

An increase in the markup factor $m \equiv \frac{\phi-1}{\phi}$, increases incumbents' and entrants' R&D, growth, and the incumbents' share of total R&D labor, and decreases inventor productivity, g/L^{RD}.

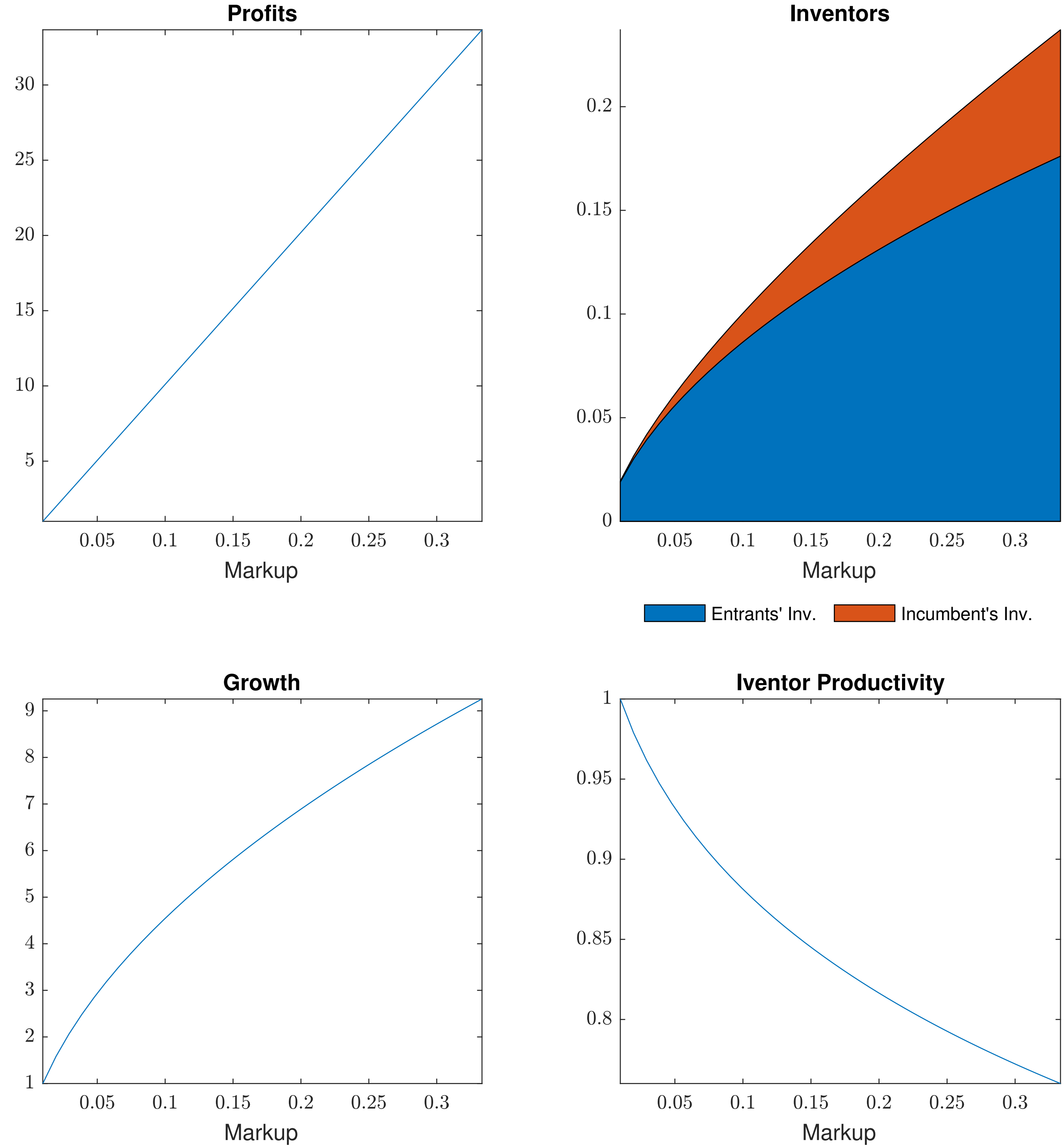
Figure 5 displays the comparative statics for various aggregates of interest in an illustrative calibration for the single-sector model, where the supply of inventors is assumed to be perfectly elastic. In this setting, an increase in the markup factor $\frac{(\phi-1)}{\phi}$:

1. Increases profits;
2. Increases both entrants' and incumbents' RD;
3. Increases the incumbents' RD share;
4. Increases sectoral growth;
5. Reduces sectoral RD productivity.

Clearly, increase in growth is a byproduct of the perfectly elastic supply of inventors, who get drawn into R&D by increased profits. It is apparent that the increased in growth is driven by the increase in entrants' R&D activity, while the decrease in inventor productivity is driven by the increase in the inventors' share hired by monopolists engaged in defensive innovation.

[9] The proof assumes $\delta = 0$ for analytical simplicity, but this assumption does not seem necessary. All the following figures were obtained from a calibration with $\delta > 0$.

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 $\frac{\phi-1}{\phi} = .01$. The calibration is chosen to highlight the properties of the model.

4.2 Two sectors

Now consider the presence of two sectors. Keep the workers' wage as given, but now assume that inventors' wages, w are determined in equilibrium. First assume that the final good is also a Cobb-Douglas with share β_i of two sectors as described above. Equilibrium in the inventor market is given by:

$$X^s(w) = \sum_{i=\{1,2\}} (RD_i(w)),$$

where X^s is the inventors' supply schedule.

4.2.1 Complete solution

Differently for before, each sector has a price index. By the above assumptions, the final good is produced according to:

$$Y = \prod Y_i^{\beta_i}. \tag{9}$$

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

$$Y_i = \beta_i \frac{Y}{P_i}. \tag{10}$$

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

$$P_i = p_{is} = wc_i, \; Y_i = \beta_i \frac{Y}{wc_i} \tag{11}$$

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:

$$\ell_{is} = \frac{c_i y_{is}}{\phi_i}, \Rightarrow L_i = \int \ell_{is} \mathrm{d}s = Y \frac{\beta_i}{\phi_i w} \tag{12}$$

Assuming a rigid production labor supply:[10]

$$L^s(w) = L = \frac{Y}{w}\left(\sum \frac{\beta_i}{\phi_i}\right). \tag{13}$$

Which gives:

$$L_i = L\frac{\frac{\beta_i}{\phi_i}}{\sum \frac{\beta_i}{\phi_i}}, Y_i = L\frac{\frac{\beta_i}{c_i}}{\sum \frac{\beta_i}{\phi_i}} \tag{14}$$

Which gives:

$$Y = L\prod_i \left(\frac{\frac{\beta_i}{c_i}}{\sum \frac{\beta_i}{\phi_i}}\right)^{\beta_i} \tag{15}$$

Thus, growth is:

$$-\sum \beta_i \Delta \log c_i \tag{16}$$

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 $\zeta w^R, \alpha_I w^R$.

4.2.2 Research Equilibrium in the two-sector model

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

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

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

$$\rho v_i^I(1) = \max_{x_I}\left(\frac{\phi_i - 1}{\phi_i}\right)\beta_i - \alpha_I w^{RD}\frac{x_I^2}{2} + x_I\left(v^I(\omega) - v^I(1)\right) - x_{e,1}\left(v^I(1)\right) \tag{17}$$
$$\rho v_i^I(\omega_i) = \left(\frac{\phi_i - 1}{\phi_i}\right)\beta_i + \delta\left(v(1) - v(\omega)\right) - x_{e,\omega}\left(v^I(\omega)\right), \tag{18}$$

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

[10] 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.

Given a normalized wage, each sector demands:

$$x_{e,\omega,i}\left(w^{RD}\right)=\frac{v_i^I(1)}{w^{RD}\omega\zeta_i}, \tag{19}$$

$$x_{I,i}\left(w^{RD}\right)=\frac{\left(v_i(\omega_i)-v_i(1)\right)}{w^{RD}\alpha_{I,i}}, \tag{20}$$

The stationary distribution within each sector is given by:

$$\mu_{\omega,i}\left(w^{RD}\right)=\frac{x_{I,i}\left(w^{RD}\right)}{x_{e,\omega,i}\left(w^{RD}\right)+\delta_i+x_{I,i}\left(w^{RD}\right)}, \tag{21}$$

$$\mu_{1,i}\left(w^{RD}\right)=\frac{x_{e,\omega,i}\left(w^{RD}\right)+\delta_i}{x_{e,\omega,i}\left(w^{RD}\right)+\delta_i+x_{I,i}\left(w^{RD}\right)}, \tag{22}$$

$$\mu_{e,1,i}\left(w^{RD}\right)=\frac{\omega_i\left(x_{e,\omega,i}\left(w^{RD}\right)+\delta\right)\mu_{1,i}+\delta_i\mu_{\omega,i}}{\left(x_{I,i}+\omega_i\left(x_{e,\omega,i}\left(w^{RD}\right)+\delta_i\right)\right)}, \tag{23}$$

$$\mu_{e,\omega,i}\left(w^{RD}\right)=\frac{\omega_i\mu_{1,i}x_{I,i}\left(w^{RD}\right)-\omega_i\delta_i\mu_{\omega,i}}{\left(x_{I,i}+\omega_i\left(x_{e,\omega,i}\left(w^{RD}\right)+\delta_i\right)\right)}+\mu_{\omega,i}. \tag{24}$$

Sector RD labor demand is given by:

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

Market clearing for inventors then reads:

$$L^{RD,s}\left(w^{RD}\right)=\sum_i\left\{\mu_{\omega,i}\left(w^{RD}\right)\left(\zeta_i\omega_i x_{e,\omega,i}\left(w^{RD}\right)\right)+\mu_{1,e,i}\left(w^{RD}\right)\zeta_i x_{e,1,i}\left(w^{RD}\right)+\mu_{1,i}\left(w^{RD}\right)\alpha_I\frac{x_{I,i}^2\left(w^{RD}\right)}{2}\right\}. \tag{25}$$

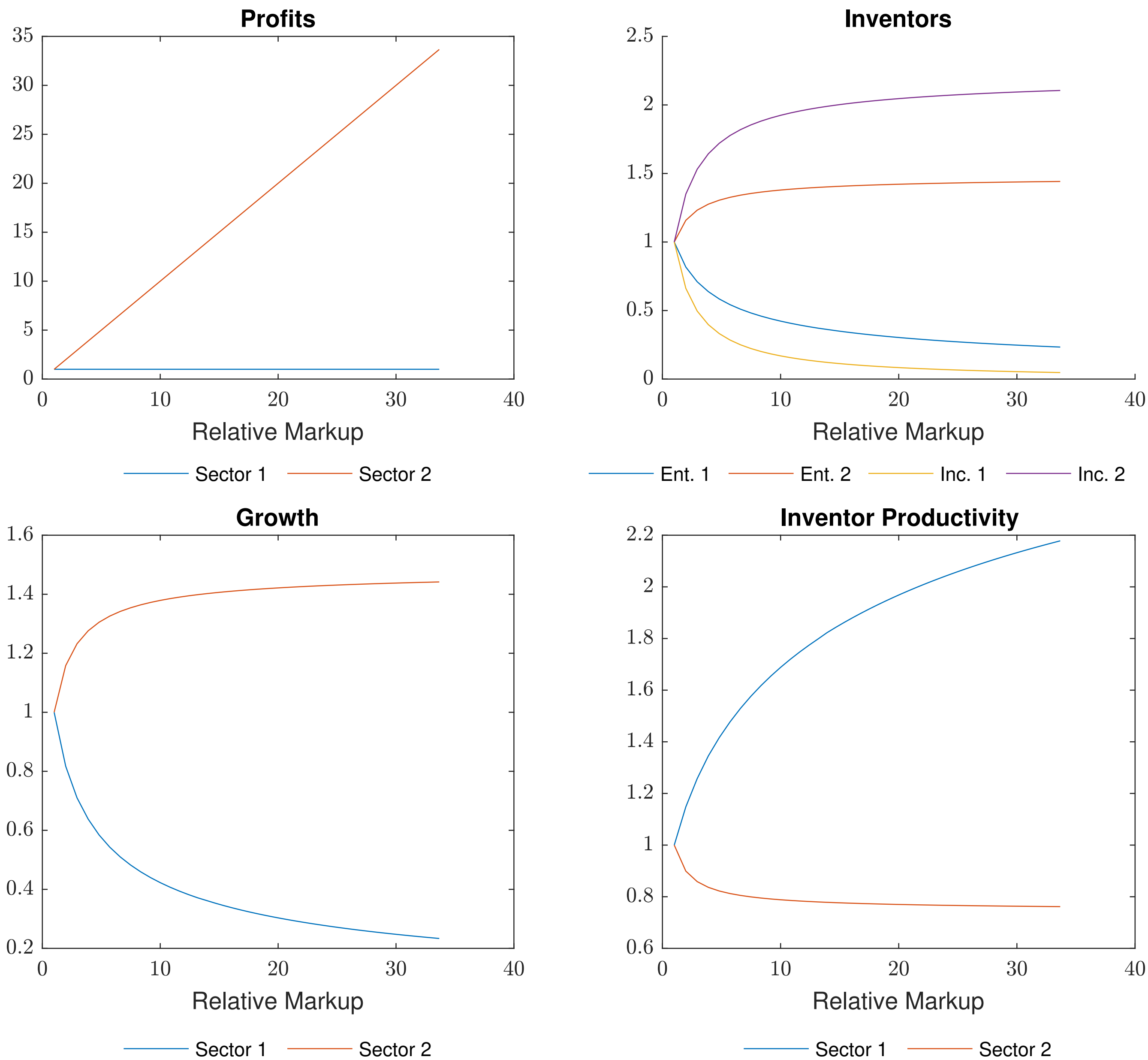
4.2.3 Comparative Statics in General Equilibrium

Figure 6 displays the comparative statics for various aggregates of interest in an illustrative calibration for the two-sector model, where the supply of inventors is set to $L_{RD}=1$, and $\beta_1=\beta_2=.5$. In this setting, an increase in the markup of sector 2 relative to sector 1:

1. Increases profits in sector 2;
2. Increases both entrants' and incumbents' RD in sector 2;
3. Increases the incumbents' R&D share in sector 2;
4. Increases sector 2's growth;
5. Reduces sector 2's R&D productivity.

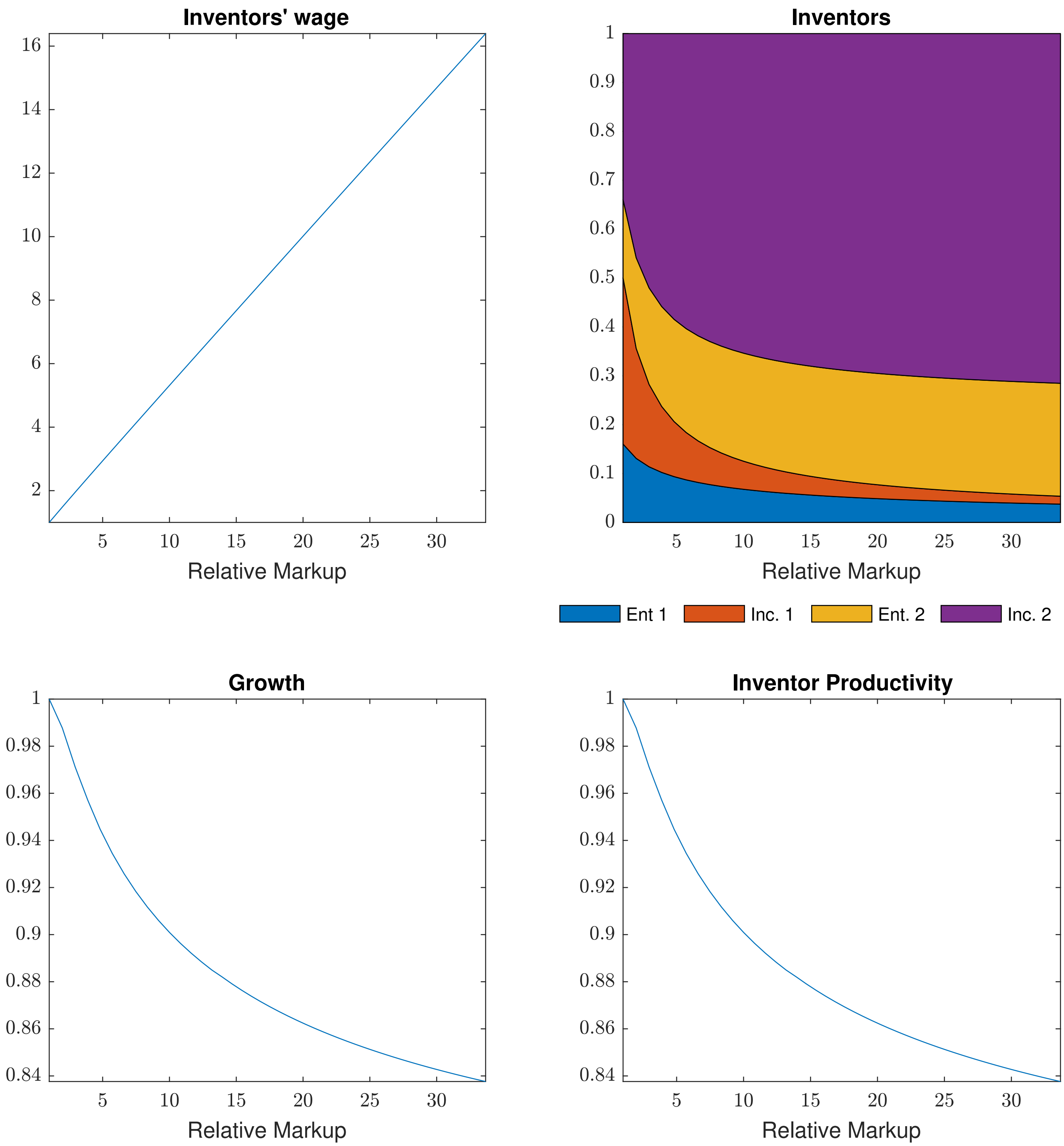
Sector 1 faces specular changes for 2.-4. Figure 7 shows the aggregate effects of the sectoral shifts presented in 6. In particular, increased markups in sector 2 bid up aggregate inventors' wages, which is accompanied by a reallocation of researchers away from sector 1 and towards sector 2. At the same time, the inventors accrue more than proportionally to incumbents in sector 2, increasing defensive research. As a result of higher defensive research, productivity falls in sector 2. The increased productivity is sector 1 is not enough to compensate this effect on aggregate output. Lower aggregate growth and productivity are also a consequence of misallocation across sectors, since in this example, where $\beta_1 = \beta_2 = .5$, aggregate growth is maximized when sectoral growth is the same for both sectors.

Figure 6: 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 calibration is chosen to highlight the properties of the model.

Figure 7: 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 calibration is chosen to highlight the properties of the model.

References

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

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.

B. Hall, A. Jaffe, and M. Trajtenberg. Market Value and Patent Citations: A First Look. Technical Report w7741, National Bureau of Economic Research, Cambridge, MA, June 2000.

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

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

A Data Construction Details

Coming soon...

B Extended Model

Suppose that incumbents can invest in innovating projects, which are successful with endogenous intensity, x_I. If the project is successful, the firm obtains a patent on the innovation it discovered. I assume that, with exogenous probability $\lambda \in [0,1]$ the innovation is compatible with existing firm's technology, and is adopted by the incumbent. This parameter can be interpreted as the probability that the realized innovation is suitable to the existing technology of the firm. An alternative interpretation of this parameter is related to incremental innovation. As λ decreases towards 0, incumbents R&D projects contribute less and less to aggregate growth. Thus, λ is positively related to how radical incumbents' projects are. If adoption occurs, the incumbent's unit costs drop to $\frac{c_t}{(1+\eta_I)\phi}$in instant t, increasing the contemporaneous technology gap between incumbents an entrants. As with entrants' innovation, I assume that if the incumbent manages to implement the innovation, remaining entrants' unit costs at instant $t+\Delta t$ fall to $\frac{c_t}{(1+\eta_I)}$, so that successful innovation does not alter the long-run technology gap between incumbents and entrants. For simplicity, I further assume that there can only be a patent barrier of size ω. Under these assumptions, the incumbents' value functions read:

$$\begin{aligned}
\rho v(1) &= \max_{x_I} \left(\frac{\phi_i - 1}{\phi_i}\right) - \alpha_I \frac{x_I^2}{2} + x_I \left[\lambda \frac{\eta_I}{\left(1+\eta_I\right)\phi_i} + v(\omega) - v(1)\right] - x_{e,1} v(1), \\
\rho v(\omega) &= \max_{x_{I,\omega}} \left(\frac{\phi_i - 1}{\phi_i}\right) - \alpha_I \frac{x_I^2}{2} + x_{I,\omega} \lambda \frac{\eta_I}{\left(1+\eta_I\right)\phi_i} + \delta\left(v(1) - v(\omega)\right) - x_{e,\omega} v(\omega), \qquad (26)
\end{aligned}$$

where the additional term, $\lambda\frac{\eta_I}{(1+\eta_I)\phi_i}$,captures the expected additional instantaneous profits accruing to incumbents from a realized innovation. Note that the LOM of the system in this case is unchanged relative to the baseline case, and so are all the comparative statics, once the appropriate parametric restrictions are in place.

C Linear Costs

Suppose that now the incumbents' research technology has:

$$\begin{aligned}
\rho v(1) &= \max_{x_I} \left(\frac{\phi_i - 1}{\phi_i}\right) - \alpha_I x_I \bar{X} + x_I \left[v(\omega) - v(1)\right] - x_{e,1} v(1), \\
\rho v(\omega) &= \max_{x_{I,\omega}} \left(\frac{\phi_i - 1}{\phi_i}\right) + \delta\left(v(1) - v(\omega)\right) - x_{e,\omega} v(\omega), \qquad (27)
\end{aligned}$$

with $\bar{X} = x_I + x_{e,1}$. This gives:

$$X = \frac{[v(\omega) - \bar{v(1)}]}{\alpha} = x_{e,1} + x_I \Rightarrow x_I = \frac{v(\omega) - v(1)}{\alpha} - \frac{v(1)}{\zeta}$$
$$\Rightarrow x_I = \frac{\zeta\, v(\omega) - (\alpha + \zeta)\, v(1)}{\alpha\zeta}$$

Now note that with these costs, in equilibrium:

$$\rho\, v(1) = m - \frac{v(1)^2}{\zeta} \tag{28}$$

Yielding:

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

The second equation gives:

$$v(\omega) = \frac{m + \delta\, v(1)}{\rho + \delta + \frac{v(1)}{\zeta\omega}}$$

Suppose $\delta = 0$:

$$v(\omega) = \frac{m\zeta\omega}{\rho\zeta\omega + v(1)}$$

And:

$$\zeta\, v(\omega) - (1 + \zeta)\, v(1) > 0$$

IFF:

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

By 28:

$$\left(1 - \frac{1+\zeta}{\zeta\omega}\right) m > \left(\rho \frac{\omega - 1}{\omega}\right) \frac{1+\zeta}{\zeta} v(1)$$
$$\left(\frac{\zeta\,(\omega - 1) - 1}{\zeta\omega}\right) m > \left(\rho \frac{\omega - 1}{\omega}\right) \frac{1+\zeta}{\zeta} \sqrt{\zeta}\sqrt{m}$$
$$(\zeta\,(\omega - 1) - 1)\, m > \left(\rho\sqrt{\zeta}\,(\omega - 1)\right)(1+\zeta)\sqrt{m}$$

Suppose $\zeta\,(\omega - 1) > 1$. Then with:

$$\sqrt{m} > \frac{\rho\sqrt{\zeta}\,(\omega - 1)\,(1+\zeta)}{\zeta\,(\omega - 1) - 1} \iff m > \left(\frac{\rho\sqrt{\zeta}\,(\omega - 1)\,(1+\zeta)}{\zeta\,(\omega - 1) - 1}\right)^2,$$

Incumbents' research is positive.

Even simpler. Suppose that both entrants and incumbents face the same crowding externalities. Then, we get:

$$\frac{v(\omega)-v(1)}{\alpha}=\bar{X}=\frac{v(1)}{\zeta} \tag{29}$$

Then:

$$v(1)=\frac{\zeta}{\alpha+\zeta}v(\omega)$$

From 26:

$$\rho\left(v(\omega)-v(1)\right)=x_{e,1}v(1)-\frac{x_{e,1}}{\omega}v(\omega)$$
$$\rho\omega\left(v(\omega)-v(1)\right)=x_{e,1}\left(\omega v(1)-v(\omega)\right)$$

Since from above $v(1)<v(\omega)$, we need to assume

$$\frac{\zeta\omega}{\alpha+\zeta}-1>0$$

to avoid contradiction, that is:

$$\zeta\left(\omega-1\right)>\alpha$$

Under this assumption:

$$\rho\omega\left(\frac{\alpha+\zeta}{\zeta}v(1)-v(1)\right)=x_{e,1}\left(\omega v(1)-\frac{\alpha+\zeta}{\zeta}v(1)\right)$$
$$\frac{\rho\omega\alpha}{\zeta\left(\omega-1\right)-\alpha}=x_{e,1}$$

From:

$$\begin{aligned}
x_I&=\frac{v(1)}{\zeta}-x_{e,1}\\
x_I&=\frac{v(1)}{\zeta}-\frac{\rho\omega\alpha}{\zeta\left(\omega-1\right)-\alpha}\\
x_I&=\frac{-\rho\zeta+\sqrt{\left(\rho\zeta\right)^2+4\zeta m}}{2\zeta}-\frac{\rho\omega\alpha}{\zeta\left(\omega-1\right)-\alpha}\\
&>\frac{\sqrt{m}}{\sqrt{\zeta}}-\frac{\rho\omega\alpha}{\zeta\left(\omega-1\right)-\alpha}\\
&>\frac{\left(\zeta\left(\omega-1\right)-\alpha\right)\sqrt{m}-\sqrt{z}\rho\omega\alpha}{\sqrt{\zeta}}>0\\
&\iff m>\zeta\left(\frac{\rho\omega\alpha}{\zeta\left(\omega-1\right)-\alpha}\right)^2
\end{aligned}$$

Is there a contradiction with $m < 1$? To avoid that we need:

$$\zeta(\omega - 1) > \left(\sqrt{\zeta}\rho\omega + 1\right)\alpha$$

Suppose $\zeta < 1$,then:

$$\zeta\left(\omega\left(1 - \rho\alpha\right) - 1\right) > \alpha,$$

which is feasible. Note that throughout there are no real restrictions on ω. Thus, in the linear model, entrants do not increase their investments with markups, while incumbents do, which directly leads to the main result in the paper. In any event, it is true that starting from an equilibrium where incumbents do invest in research, their investment increases in the markup, while the ones from entrants in unprotected sectors does not. This is because increased investment by incumbents increases aggregate costs enough to make additional investment unattractive to entrants.

Note that there is only entrants' research in protected sectors. Given the crowding externalities specified above, we have:

$$x_{e,\omega} = \frac{v(1)}{\zeta\omega}.$$

With $\omega > 1$:

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

Note that growth in this context is given by:

$$\eta\left(\mu_{e,1}x_{e,1} + \mu_{e,\omega}x_{e,\omega}\right)$$

and inventors' productivity is:

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

The incumbents' share is proportional to:

$$\text{Sh.Inc} \propto \frac{\mu_1 x_I}{\mu_{e,1}x_{e,1} + \omega\mu_{e,\omega}x_{e,\omega}}$$

Below we found:

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

$$\mu_{e,1} = \frac{x_{e,1}\mu_1}{x_{e,1} + x_I}$$

Note that:

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

Thus:

$$
\begin{aligned}
\mu_{e,\omega} &= 1 + \frac{x_{e,1}\mu_1\left(x_I - x_{e,\omega}\right) - \left(x_I\right) x_{e,\omega}\mu_1}{x_{e,\omega}\left(x_{e,1} + x_I\right)}, \\
\mu_{e,\omega} &= 1 + \frac{\left(\omega x_{e,\omega} - x_I\right)\left(x_I - x_{e,\omega}\right) - x_I x_{e,\omega}}{x_{e,\omega}\left(\omega x_{e,\omega}\right)}\mu_1, \\
&= 1 + \frac{\omega x_{e,\omega} - \omega x_{e,\omega}^2 - x_I^2}{x_{e,\omega}\left(\omega x_{e,\omega}\right)}\mu_1 \\
&= 1 + \frac{1 - x_{e,\omega} - \frac{x_I^2}{\omega x_{e,\omega}}}{x_I + x_{e,\omega}} \\
&= \frac{1 + x_I\left(1 - \frac{x_I}{\omega x_{e,\omega}}\right)}{x_I + x_{e,\omega}} \\
&= \frac{1 + x_{e,1}\frac{x_I}{\omega x_{e,\omega}}}{x_I + x_{e,\omega}}
\end{aligned}
$$

And:

$$
\begin{aligned}
\mu_{e,1} &= \frac{x_{e,1}\mu_1}{x_{e,1} + x_I} \\
\mu_{e,1} &= \frac{\omega x_{e,\omega} - x_I}{\omega\left(x_I + x_{e,\omega}\right)}
\end{aligned}
$$

Thus:

$$
\begin{aligned}
\mu_{e,1}x_{e,1} + \omega\mu_{e,\omega}x_{e,\omega} &= \\
\frac{\omega x_{e,\omega} - x_I}{\omega\left(x_I + x_{e,\omega}\right)}x_{e,1} + \omega\frac{1 + x_{e,1}\frac{x_I}{\omega x_{e,\omega}}}{x_I + x_{e,\omega}}x_{e,\omega} &= \\
\frac{\left(\omega x_{e,\omega} - x_I\right)^2 + \omega^2 x_{e,\omega} + \omega\left(\omega x_{e,\omega} - x_I\right)x_I}{\omega\left(x_I + x_{e,\omega}\right)}
\end{aligned}
$$

Thus:

$$
\begin{aligned}
\text{Sh.Inc} \propto &\ \frac{\mu_1\left(x_{e,1} + x_I\right)}{\frac{\left(\omega x_{e,\omega} - x_I\right)^2 + \omega^2 x_{e,\omega} + \omega\left(\omega x_{e,\omega} - x_I\right)x_I}{\omega\left(x_I + x_{e,\omega}\right)}} \\
&= \frac{x_{e,\omega}^2}{\left(\omega x_{e,\omega} - x_I\right)^2 + 2\omega^2 x_{e,\omega} - \omega x_I^2} \\
&= \frac{x_{e,\omega}^2}{\omega^2 x_{e,\omega}^2 - 2\omega x_{e,\omega}x_I + x_I^2 - \omega x_I^2} \\
&= \frac{1}{\omega^2 - 2\omega\frac{x_I}{x_{e,\omega}} - (\omega - 1)\left(\frac{x_I}{x_{e,\omega}}\right)^2},
\end{aligned}
$$

which is unambiguously increasing in $\frac{x_I}{x_{e,\omega}}$ since $\omega > 1$. Therefore, the linear model also predicts an increase in the incumbents' share of inventors following an increase in the markup, and the comparative statics are qualitatively the same as in Proposition **??**.

D Omitted Proofs and Derivations

D.1 One-sector model

Proof of Lemma 4.1. The values, $v(1), v(\omega)$, satisfy the system:

$$\rho v(1) - m_i - \mathbf{1}\{v(\omega) - v(1) > 0\} \frac{(v(\omega) - v(1))^2}{2\alpha_I} + \frac{(v(1))^2}{\zeta} = 0 \tag{30}$$

$$\rho v(\omega) - m_i - \delta (v(1) - v(\omega)) + \frac{v(1)}{\zeta\omega} (v(\omega)) = 0 \tag{31}$$

Then:

$$\rho (v(\omega) - v(1)) = \delta (v(1) - v(\omega)) - \frac{v(1)}{\zeta\omega} (v(\omega)) - \frac{(v(\omega) - v(1))^2}{2\alpha_I} + \frac{v(1)}{\zeta} v(1)$$

$$\left(\rho + \delta + \mathbf{1}\{v(\omega) - v(1) > 0\} \frac{(v(\omega) - v(1))}{2\alpha_I}\right) (v(\omega) - v(1)) = \frac{v(1)}{\zeta} \left(v(1) - \frac{v(\omega)}{\omega}\right)$$

Suppose that $v(\omega) < v(1)$. This implies that the left hand side of the above expression is strictly smaller than 0, while the right hand side is strictly positive since $\omega > 1$. 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, $\omega v(1) > v(\omega) > v(1)$.

Total differentiation of the system, (30)(31), gives:

$$\underbrace{\begin{bmatrix} \rho + \frac{v(\omega)-v(1)}{\alpha_I} + 2\frac{v(1)}{\zeta} & -\frac{v(\omega)-v(1)}{\alpha_I} \\ -\delta + \frac{v(\omega)}{\zeta\omega} & \rho + \delta + \frac{v(1)}{\zeta\omega} \end{bmatrix}}_{\equiv J} \begin{bmatrix} \mathrm{d}v(1) \\ \mathrm{d}v(\omega) \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathrm{d}m_i = 0 \tag{32}$$

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 (32) gives:

$$\begin{bmatrix} \frac{\mathrm{d}v(1)}{\mathrm{d}m_i} \\ \frac{\mathrm{d}v(\omega)}{\mathrm{d}m_i} \end{bmatrix} = \frac{1}{\det J} \begin{bmatrix} \frac{\rho\zeta\omega + v(1)}{\zeta\omega} + \delta & \frac{v(\omega)-v(1)}{\alpha_I} \\ \delta - \frac{v(\omega)}{\zeta\omega} & \rho + \frac{v(\omega)-v(1)}{\alpha_I} + 2\frac{v(1)}{\zeta} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Since the first row is strictly positive,

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

Subtracting line by line gives:

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

since $\omega > 1$ and $\omega v(1) > v(\omega)$. It follows that:

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

□

Derivation of the Stationary Distribution The law of motion of the distribution across states satisfies:

$$\dot{\mu}_1 = -\left(x_I + x_{e,1}\right)\mu_1 + \delta\mu_\omega + x_{e,\omega}\mu_{e,\omega} + x_{e,1}\mu_{e,1}, \tag{33}$$

$$\dot{\mu}_\omega = -\left(x_{e,\omega} + \delta\right)\mu_\omega + x_I\mu_1, \tag{34}$$

$$\dot{\mu}_{e,1} = -\left(x_{e,1} + x_I\right)\mu_{e,1} + x_{e,1}\mu_1 + \delta\mu_{e,\omega}, \tag{35}$$

$$\dot{\mu}_{e,\omega} = -\left(x_{e,\omega} + \delta\right)\mu_{e,\omega} + x_{e,\omega}\mu_\omega + x_I\mu_{e,1}, \tag{36}$$

By equation (34):

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

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

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

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

$$\begin{bmatrix} -\delta & x_{e,1} + x_I \\ x_{e,\omega} + \delta & -x_I \end{bmatrix}\begin{bmatrix} \mu_{e,\omega} \\ \mu_{e,1} \end{bmatrix} = \begin{bmatrix} x_{e,1}\mu_1 \\ x_{e,\omega}\mu_\omega \end{bmatrix}. \tag{37}$$

Solving the system (37):

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

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

By the optimal solution for entrants:

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

so (37) is solved for:

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

Proof of Proposition 4.2. The increase in R&D and growth descend directly from Lemma 4.1. Note that the incumbents' share of researchers increases with:

$$\frac{\alpha_I}{\zeta}\frac{x_I}{2}\frac{\mu_1 x_I}{\mu_{e,1}x_{e,1} + \mu_{e,\omega}\omega x_{e,\omega}} = \frac{\mu_1 x_I}{\omega x_{e,\omega}\left[\mu_{e,1} + \mu_{e,\omega}\right]}$$

Now note:

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

with $\delta = 0$:

$$\begin{aligned}
\mu_{e,1} + \mu_{e,\omega} &= \frac{\omega x_{e,\omega}\mu_1}{\omega x_{e,\omega} + x_I} + \frac{\omega x_I \mu_1 + \left(\omega x_{e,\omega} + x_I\right)\mu_\omega}{\omega x_{e,\omega} + x_I} \\
&= \frac{\omega x_{e,\omega}\mu_1}{\omega x_{e,\omega} + x_I} + \frac{\omega x_I \mu_1 + \left(\omega x_{e,\omega} + x_I\right)\left(1 - \mu_1\right)}{\omega x_{e,\omega} + x_I} \\
&= (\omega - 1)\,\mu_1 \frac{x_I}{\omega x_{e,\omega} + x_I} + 1
\end{aligned}$$

Thus the above becomes:

$$\begin{aligned}
\frac{\alpha_I}{\zeta\omega}\frac{x_I}{2}\frac{\mu_1 x_I}{x_{e,\omega}\left[\mu_{e,1} + \mu_{e,\omega}\right]} &= \frac{\alpha_I}{\zeta\omega}\frac{x_I}{2}\frac{\mu_1 x_I}{x_{e,\omega}\left[1 + (\omega - 1)\,\mu_1 \frac{x_I}{\omega x_{e,\omega} + x_I}\right]} \\
&= \frac{\alpha_I}{\zeta\omega}\frac{x_I}{2}\frac{x_I}{\frac{x_{e,\omega}}{\mu_1} + (\omega - 1)\,x_{e,\omega}\frac{x_I}{\omega x_{e,\omega} + x_I}} \\
&= \frac{\alpha_I}{\zeta\omega}\frac{x_I}{2}\frac{x_I\left(\omega x_{e,\omega} + x_I\right)}{\left(x_I + x_{e,\omega}\right)\left(\omega x_{e,\omega} + x_I\right) + (\omega - 1)\,x_{e,\omega}x_I} \\
&= \frac{\alpha_I}{\zeta\omega}\frac{x_I}{2}\frac{\left(\omega + \frac{x_I}{x_{e,\omega}}\right)}{\left(1 + \frac{x_{e,\omega}}{x_I}\right)\left(\omega + \frac{x_I}{x_{e,\omega}}\right) + (\omega - 1)} \\
&= \frac{\alpha_I}{\zeta\omega}\frac{x_I}{2}\frac{(\omega + z)}{\left(1 + \frac{1}{z}\right)(\omega + z) + (\omega - 1)},
\end{aligned}$$

where I define $z \equiv x_I / x_{e,\omega}$.Above, I showed that $\frac{\partial x_I}{\partial m} > 0$. To show that the fraction is also increasing , note that:

$$\frac{\partial\left[\frac{(\omega+z)}{\left(1+\frac{1}{z}\right)(\omega+z)+(\omega-1)}\right]}{\partial z}=\frac{\left(1+\frac{1}{z}\right)(\omega+z)+(\omega-1)-(\omega+z)\left[-\frac{1}{z^2}(\omega+z)+\left(1+\frac{1}{z}\right)\right]}{\left[\left(1+\frac{1}{z}\right)(\omega+z)+(\omega-1)\right]}$$
$$=\frac{\frac{\omega+z}{z^2}(\omega+z)+(\omega-1)}{\left[\left(1+\frac{1}{z}\right)(\omega+z)+(\omega-1)\right]}>0$$

Thus the sign of the fraction only depends on the sign of the derivative of z with respect to m.

By definition:

$$z\equiv\frac{\zeta\omega}{\alpha_I}\left(\frac{v(\omega)}{v(1)}-1\right)m.$$

As shown above, positive research implies that $\omega v(1)>v(\omega)$, or equivalently that $\frac{v(\omega)}{v(1)}<\omega$. This gives

$$z\equiv\frac{x_I}{x_e}<\frac{\zeta\omega(\omega-1)}{\alpha_I}\Rightarrow x_I<\frac{\zeta\omega(\omega-1)}{\alpha_I}x_e.$$

Further:

$$\operatorname{sign}\left(\frac{\partial z}{\partial m}\right)=\operatorname{sign}\left(\frac{\partial\left(v(\omega)/v(1)\right)}{\partial m}\right)=\operatorname{sign}\left(\frac{\partial v(\omega)}{\partial m}v(1)-\frac{\partial v(1)}{\partial m}v(\omega)\right). \tag{38}$$

By Lemma 4.1:

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

Thus (38) has the same sign as:

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

Since there is positive research, it holds:

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

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

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

Note that, by definition of z:

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

By the 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. This gives:

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

Therefore:

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

by the main assumption in the statement. The decrease in inventor's productivity follows immediately noting that:

$$
\begin{aligned}
\frac{g}{L^{RD}} &= \eta\frac{x_{e,\omega}\left(\mu_{e,\omega} + \omega\mu_{e,1}\right)}{\zeta\omega x_{e,\omega}\left(\mu_{e,\omega} + \mu_{e,1}\right) + \alpha_I\frac{x_I^2}{2}\mu_1} \\
&= \eta\frac{x_{e,\omega}\left(\mu_{e,\omega} + \omega\mu_{e,1}\right)}{L_e\left(1 + \frac{L_I}{L_e}\right) + \alpha_I\frac{x_I^2}{2}\mu_1} \\
&= \frac{\eta}{\zeta\omega}\frac{\mu_{e,\omega} + \omega\mu_{e,1}}{\mu_{e,\omega} + \mu_{e,1}}\frac{1}{\left(1 + \frac{L_I}{L_e}\right)} \\
&= \frac{\eta}{\zeta\omega}\frac{\mu_{e,\omega} + \omega\mu_{e,1}}{(\omega-1)\mu_1\frac{x_I}{\omega x_{e,\omega} + x_I} + 1}\frac{1}{\left(1 + \frac{L_I}{L_e}\right)}
\end{aligned}
$$

Recall:

$$
\begin{aligned}
\mu_{e,1} + \omega\mu_{e,\omega} &= \frac{\omega x_{e,\omega}\mu_1}{\omega x_{e,\omega} + x_I} + \omega\frac{\omega x_I\mu_1 + \left(\omega x_{e,\omega} + x_I\right)\mu_\omega}{\omega x_{e,\omega} + x_I} \\
&= 1 - \mu_1(\omega-1)\frac{x_{e,\omega} - x_I}{\omega x_{e,\omega} + x_I}
\end{aligned}
$$

And:

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

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

$$= \frac{\omega + \omega\left[\omega\frac{x_I}{x_e} + \left(\omega + \frac{x_I}{x_e}\right)\frac{x_I}{x_e}\right]}{\omega + \omega\frac{x_I}{x_e} + \left(\omega + \frac{x_I}{x_e}\right)\frac{x_I}{x_e}}$$

$$\equiv \omega\frac{1+g\left(\frac{x_I}{x_e}\right)}{\omega + g\left(\frac{x_I}{x_e}\right)}$$

Thus the ratio is decreasing in $g(\cdot)$ which is itself increasing in $x_I/x_{e,\omega}$. By the previous points, increases in markup lower growth by concentrating more resources with incumbents. □

D.2 Two-Sector Model

E Additional Results and Robustness

E.1 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 1 and 2 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 9: Regressions of Change in 4-digit Knowledge Market Share of Total Inventors over Change in HHI Measures, Long-Differences, 1997-2012

	Ch. 4d K.M. Eff. Inv. Share (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Ch. HHI lower bound	74.172+		74.814+		74.177+	
	(40.957)		(41.208)		(41.047)	
Ch. HHI		71.749**		71.749**		71.583**
		(24.464)		(24.464)		(24.433)
4D Knowledge Market						
Sample	Full Sample	Full Sample	Trim Outliers	Trim Outliers	Mahalanobis 5%	Mahalanobis 5%
Weight	Sales	Sales	Sales	Sales	Sales	Sales
Observations	157	80	156	80	150	72

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), 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 10: 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

	Ch. 4d K.M. Eff. Inv. Share (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Ch. HHI lower bound	71.724+	67.160+	72.123+	67.736+	71.772+	68.398+
	(39.265)	(37.176)	(39.530)	(37.504)	(39.316)	(37.717)
Ch. Log Real Sales	1.864*	1.422*	1.852*	1.402+	1.878*	1.443+
	(0.766)	(0.717)	(0.764)	(0.712)	(0.774)	(0.745)
4D 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	156	155	150	142

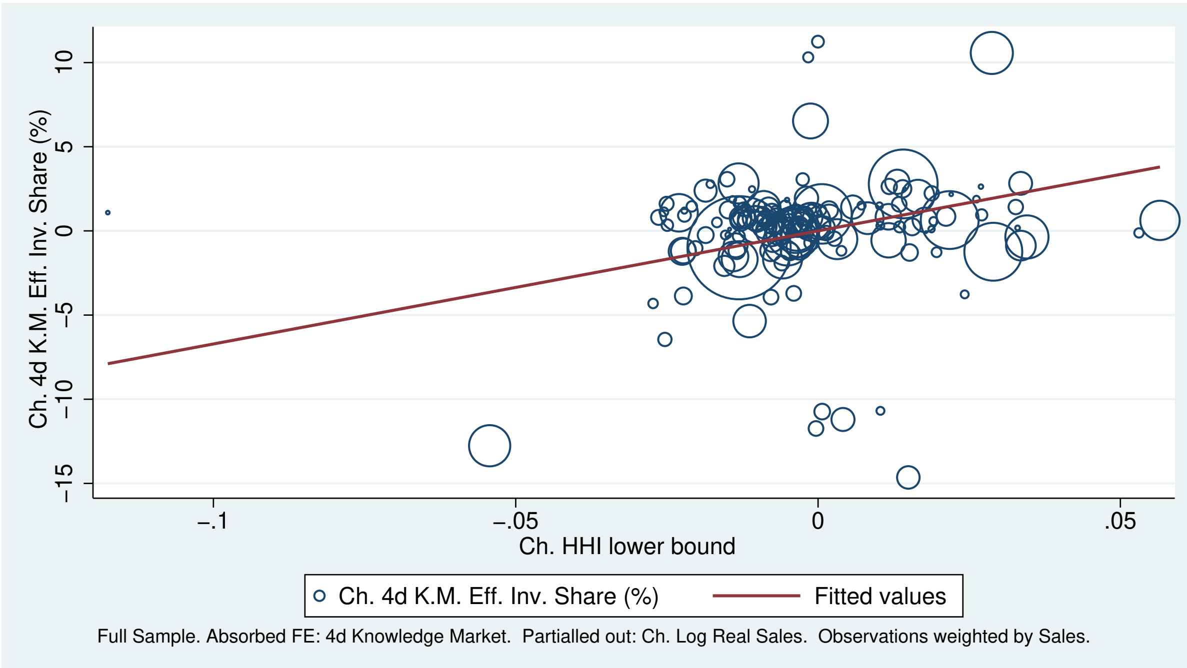
(b) Controlling for Change in Log Real Sales per Company

	Ch. 4d K.M. Eff. Inv. Share (%)					
	(1)	(2)	(3)	(4)	(5)	(6)
Ch. HHI lower bound	35.230**	20.783+	35.230**	20.783+	35.154**	22.854*
	(12.759)	(10.615)	(12.759)	(10.615)	(12.647)	(11.197)
Ch. Log Real Sales per company	0.175	-0.040	0.175	-0.040	0.300	-0.055
	(0.382)	(0.253)	(0.382)	(0.253)	(0.460)	(0.346)
4D 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

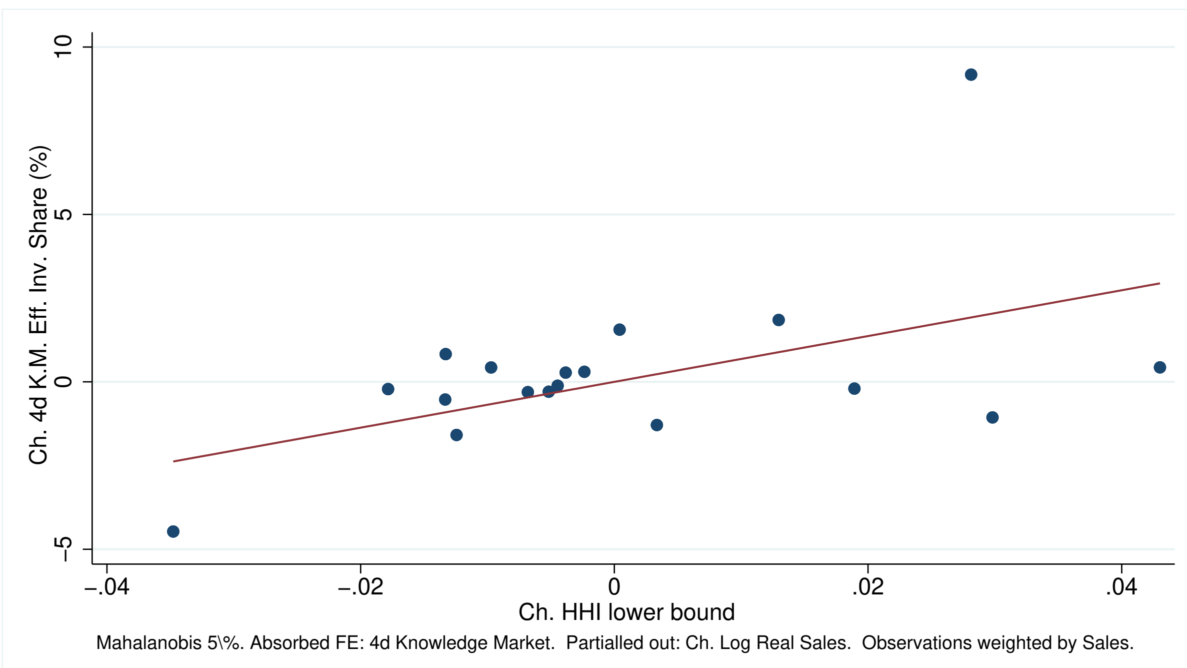
Note: Regressions weighted by sales in 2012; Robust standard errors in parentheses; Symbols denote significance levels $\left(+\ p<0.1,^{*}\ p<0.05,^{**}\ p<.01,^{***}\ p<.001\right)$; Checkmarks indicate the inclusion of fixed effects. This Tables presents the results of specifications (2) and (3), 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 10, 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 10. 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 10.

E.2 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:

$$\text{Lerner}_{jt} = \frac{\text{vship}_{jt} - \text{pay}_{jt} - \text{matcost}_{jt} - \text{energy}_{jt}}{\text{vship}_{jt}}, \tag{39}$$

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 11, 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 12 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 11 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 2.

Table 11: Regressions of Changes in the Lerner Index over Changes in the HHI Lower Bound, Long-Difference, 1997-2012

	Markup Change 1997-2012, 6d Lerner Index	Markup Change 1997-2012, 4d Lerner
	(1)	(2)
HHI Change 1997-2012	1.490***	1.652***
	(0.229)	(0.257)
Observations	258	258
R-squared	.1424476	.139197

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 (39)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 12: Regressions of Changes in Inventors' Share over Changes in Actual and Fitted Lerner Index, Long-Difference, 1997-2012

	Ch. 4d K.M. Eff. Inv. Share (%)	
	(1)	(2)
Markup Change 1997-2012, 4d Lerner Index	0.556	
	(5.465)	
Fitted Lerner Change		26.736*
		(13.363)
4D Knowledge Market		
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 11, and extended to all available sectors in the main sample.