Market for Patents, Monopoly, and Misallocation*

Shang-Jin Wei †, Jianhuan Xu ‡, Ge Yin§, and Xiaobo Zhang¶

March 20, 2024

Abstract

The paper studies a possible "dark side" of patent trade in enhancing the market power of monopolists. We explore the different effects of China's 2008 tax reform on patent innovations and sales across industries. In particular, although easier patent trade leads to more patent creation, the new patents are disproportionately connected to existing monopolists and are more likely to be acquired by them. Using an endogenous growth model with patent trade, we show that subsidizing patent trade could skew investors' research to appeal to the monopolists, increase the latter's monopoly power, and reduce social welfare. An optimal subsidy policy for patent trade should be contingent on the initial level of misallocation in the invention market.

*We thank seminar participants in Singapore Management University, Peking University, Tsinghua University, Wuhan University, SHUFE, and the University of Toronto for their useful comments.

†Columbia University, FISF, NBER, and CEPR

‡Singapore Management University

§University of International Business and Economics

¶Peking University and IFPRI

1 Introduction

The idea that trade can improve welfare is a foundational concept in economics. The literature on intellectual property rights similarly suggests that the trade of patents has the potential to bring about two key benefits: stronger incentives for innovation and improved allocation of innovations (see, e.g. Gans and Stern (2000), Akcigit et al. (2016), and Serrano (2018)). As a result, many governments opt to subsidize patent transfer, leading to an enormous boom in the patent market.[1]

However, despite this classical view, there is also a potential downside to the trade in patents: when it is driven by a desire to maintain monopoly power for big firms, it can deter the adoption of new technology, increase market concentration, and lead to a decline in welfare. For example, in the pharmaceutical industry, drugs that have been acquired have a higher likelihood of being discontinued than those that have not been acquired. As another example, nearly 80% of traded software and patents in Silicon Valley go to a small number of big tech giants, enhancing their already considerable market power.

This paper investigates the connection between patent trade and market power, and its impact on long-term economic growth using the lens of an endogenous growth model. The paper explores a policy shock on the patent trade market from China's reform of corporate income tax in 2008. Under Chinese law, the profits from the sale of a patent are taxed as either individual income tax or firm income tax, depending on the seller. The 2008 reform cuts the domestic-owned firm income tax rate. As a result, industries with a higher proportion of domestic corporate-owned patents experience greater tax reductions upon selling patents following the reform.[2] As the proportion of corporate-owned patents is likely to be determined by exogenous industry characteristics, such as the initial fixed costs of innovation, the various exposure to the 2008 tax reduction across industries provides an identification to isolate the true impact of patent trade on the macro economy.[3]

The study begins by identifying and documenting four distinct empirical trends. The initial observation highlights a discernible escalation in patent creation and commerce within industries possessing a substantial share of corporate-owned patents since the year 2008. This surge is predominantly driven by corporate entities, which are responsible for the majority of patent generation and sales, rather than by individual inventors. Such a pattern underscores the significant influence of the tax reform on the landscape of patent production and commercialization. Our

[1]For instance, the newly granted patents granted in China, one of the major source countries for new patents, grew with an annual rate of around 10% in 2021, while the growth rate of patent transfer is about 15%.

[2]As it will be clear later, this paper only focuses on domestic-owned patents, which account for 97% of total patents in China. Moreover, in this paper, we use "firm" and "corporate" interchangeably and do not differentiate between them.

[3]We empirically confirm that industries with substantial initial fixed costs, such as the chemical industry, consistently exhibit a higher prevalence of patents owned by corporations rather than individuals.

analysis indicates that variations in tax reform exposure can elucidate approximately half of the differences in patent activity observed across various industries before and after the pivotal year of 2008.

The second fact we unravel is the potential of tax reform to pivot the direction of research and innovation. By examining the backward citation data, we can assess the proximity of each new patent to the existing technology of incumbent firms. Post-reform, a marked trend emerges where new patents tend to more closely align with the technologies of established, particularly monopolistic, firms. This trend is notably more pronounced for patents originating from corporate innovation as opposed to those from individual inventors. In quantifiable terms, the reform measures account for 25% of the uptick in the alignment between new patents and the pre-existing patent portfolios of dominant market players.[4]

The third fact delves into the dynamics of patent transactions. There is a noticeable trend of increased patent acquisitions by large firms, particularly in industries where corporate-owned patents are predominant. In addition, firms are more likely to buy patents with a closer connection with their existing technology.[5] Therefore, the second and third facts together suggest that, within the burgeoning market for patents, innovators may skew their research efforts to appeal to big firms. This not only distorts the natural trajectory of innovation but also enhances the monopoly power within the product market. To the best of our knowledge, our study is the first to document this pattern and shed light on this mechanism in the existing literature.

In the last fact, we are interested in the causal impacts of the patent trade on the industry's aggregate productivity and monopoly power. To address the endogeneity, we instrument the patent trade across industries using the various exposures to the trade tax reduction in the reform in a similar spirit as the Bartik instruments.[6] We find that the increase in trade can raise the industrial average markup more pronounced than the productivity: a 1% increase in patent trade increases the markup two times greater. In other words, although the patent trade can bring improvement in productivity, it reduces the competition and increases the monopoly power at the same time.[7] Determining if the rise in patent trade is advantageous to the overall economy requires a nuanced quantitative evaluation of its benefits and drawbacks.

[4]Bryan and Hovenkamp (2020) use a very different model from ours but theoretically obtain a similar point: the acquisition of startups can bias the research direction of startups towards industry leaders. However, as far as we know, we are the first to document this pattern empirically.

[5]This fact is consistent with the finding by Akcigit et al. (2016), which highlights the complementarity between buyers' existing patents and newly acquired patents.

[6]The first stage result indicates that those industries that enjoy a 1% more patent trade tax cut in the reform due to a high corporate-owned patents share show a 0.3% more increase in patent trade.

[7]In a perfectly competitive market, the productivity change and the markup change should be the same (although in the opposite direction).

Inspired by these facts, we build an endogenous growth model to quantify the welfare impact of the pro-patent trade policy. The model introduces the patent trade market into the Schumpeterian creative-destruction growth model (Klette and Kortum (2004)), in which endogenous innovation determines productivity evolution. Due to the Bertrand competition in the product market, the frontier patent holder produces as a monopolist and charges a limit price proportional to the productivity gap to the second-best patent. However, other innovators continue to bring new patents to challenge the current monopolist's leading position. To maintain its monopoly power, the monopolist relies on either internal innovation to raise its productivity or external acquisition to reduce competition from potential competitors. Both can increase the productivity gap between the current monopolist and other competitors, hence raising its markup. Therefore, markup for a given product is increasing as long as the same firm produces the product. Once a new firm successfully gets a new frontier patent and undermines the monopolist's advantage, it will become the new leader for this product and the markup tends to decline as competition intensifies. The new innovations from challengers and the monopoly power maintaining investment, either in-house innovation or acquisition, from the current monopolist, determine the equilibrium dynamics. The aggregate welfare of the economy depends on the new innovation intensities and the distortion from the monopoly power.

The model can be solved analytically and is novel in two features. First, the model introduces the patent trade market, a fast-growing market, into the growth model.[8] Given the significant impact on the productivity of the patent trade, considering it is quantitatively important. More importantly, the model can help identify two types of patent trade: "monopoly-power-driven" and "productivity-driven" acquisitions. In the former case, the current monopolist purchases the patent, while in the latter case, the firm that can best commercialize the patent is the buyer. The relative share of these two types of acquisitions determines the gain from trade. When the "monopoly-driven" acquisition becomes more frequent, it brings more distortion due to higher monopoly power, while dampening the productivity gain from trade. Hence, from the lens of our model, we can quantify both positive and negative welfare impacts from the patent trade market.

Second, the model also features how the patent trade market can change an innovator's decision. In the case in which the current monopolist has a large advantage over competitors, it is more willing to offer a higher price to acquire patents to maintain its leading position. When the trade tax is lower, it becomes more attractive to sell patents to the current monopolist. Given that the monopolist is more likely to acquire patents with a close connection, innovators will

[8] As far as we know, there is only one paper, Akcigit et al. (2016), which studies the patent trade from a growth model. However, the trade in their paper is purely driven by productivity concerns.

create new patents more appealing to the monopolist, as observed in the data. This generates a self-amplifying loop: new patents are easier to sell to big incumbents due to a tie-connection, and incumbents tend to accumulate higher monopoly powers, resulting in stronger incentives for new innovators to change their innovation directions, hence more inefficiency.[9]

Our model generates several testable implications regarding the distribution of firm size and markup. Specifically, the stationary distributions of both firm size and markup have fat tails, which are determined by the ratio of "monopoly-power-driven" trade to "productivity-driven" trade. If tax reductions incentivize inventors to innovate patents closely connected to monopolists, this ratio increases, resulting in stronger monopoly power and an increase in the tails of firm size and markups.

Moreover, the model predicts changes in buyer markup following a firm's acquisition of a patent. For instance, a "productivity-driven" acquisition tends to decrease the firm's markup, particularly among high-markup firms, because the acquisition introduces a new low-markup product to the firm's high-markup product portfolio, resulting in a larger decline in markup. As tax reform discourages "productivity-driven" acquisition, high-markup firms are more likely to experience an increase in markup than other firms following an acquisition. We utilized these predictions to validate our model.

To better understand the tax reduction in China, we calibrate the model to the data. The innovation rates of new patents and the changes in markup and firm size help us identify the model. We find that a flat trade tax reduction in 2008 significantly increases the monopoly power, but only increases the growth rate of productivity a little. The reason is simple: the tax introduces more tie-connected patents, which are mainly acquired by monopolists and are not utilized efficiently. Overall, the tax reduction hurts the aggregate welfare by -3%, since the distortion due to higher monopoly power overweights the gain from productivity improvement.

More importantly, most of the welfare loss associated with the patent trade market stems from inventors altering their research directions to maintain a closer connection with monopolists. This results in a significant increase in the share of "monopoly-power-driven" transactions within the overall patent trade market. As far as we know, our study is the first to emphasize that this channel has a quantitatively significant macroeconomic impact on the patent trade market.

We explore several counterfactual scenarios. First, we consider the optimal flat trade tax reduction for patents. Our analysis indicates that the best trade tax reduction for patents is approximately 2%, which represents only 25% of the current tax cut. However, the optimal flat

[9]This trend becomes significant in the internet industry, in which startups are more and more likely to be acquired by big incumbents.

trade tax rate is contingent upon the initial misallocation of the invention market. If the inventor's cost to commercialize the new invention is twice as high as our estimate, the optimal flat trade tax reduction is closer to 8%, which is the actual reduction that occurred in 2008. Second, we investigate the optimal patent trade tax rate, which is dependent on firm size. Our findings suggest that the optimal tax rate would be quite progressive.

Our paper makes several contributions to the literature. First, it builds on models of endogenous growth in the Schumpeterian growth model of Klette and Kortum (2004). This framework is analytically attractive and is widely used. It has been used to study industrial policies (Acemoglu et al. (2018)), to analyze the optimal protection of patents (Acemoglu and Akcigit (2012)), and to quantify the misallocation of heterogeneous markups (Peters (2020)). We extend this framework to incorporate the patent trade in an analytical way. This extra margin is not only quantitatively important but also contributes a novel mechanism to the growing literature linking trends in market concentration, productivity growth, and business dynamism within endogenous growth models. Existing literature has emphasized technological changes in innovation and accumulating intangible assets (De Ridder (2019) and Olmstead-Rumsey (2019)), demographic changes (Hopenhayn et al. (2018) and Jones (2022)), and changes in long-run real interest rates (Kroen et al. (2021)). Our paper highlights that the shifts towards larger pro-patent trade subsidies can also play an important role.

Second, a growing literature focuses on how new ideas spread in an economy and the ensuing welfare consequences. Some work stresses technology diffusion via imitation (Lucas Jr and Moll (2014) and Perla and Tonetti (2014)). Particularly, König et al. (2022) apply this idea to the Chinese innovation market. Other work stresses the frictions in the technology transfer market, such as search friction (Akcigit et al. (2016)). However, the idea diffusion in all these papers is driven by "productivity". However, this paper adds another driving force for idea diffusion in the form of patent trade: maintaining high "monopoly power". We show that this incentive may be quantitatively more important.

Third, despite the benefits of patent trade mentioned before, there is a growing concern voiced in both academic and policy debates about the potentially deleterious effects of patent trade. Several reasons have been pointed out, including extracting rents through litigation due to fragment patent ownership (Bessen and Meurer (2008)) and increasing hold-up problems in the R&D investment (Sidak (2007)). Some studies have even recommended reforming the current patent trade system (see Hall and MacGarvie (2010) for a survey). Others have pointed out the inefficiency of the pro-patent trade policies, such as the patent box- a pro-innovation subsidy program that permits acquired patents to be used in qualifying for a subsidy (see Bloom et al. (2019) for

a survey). This paper quantifies the macro implications of the dark side of the patent trade from the lens of an endogenous growth model.

Lastly, the paper is related to the literature about the misallocation of heterogeneous markups. It is well known that if the relative price of factor inputs is heterogeneous, it can create a misallocation and reduce aggregate productivity (Hsieh and Klenow (2009)). One potential reason for a dispersed factor inputs price is that markups of final goods are heterogeneous. Edmond et al. (2018) and Peters (2020) quantify the misallocation from the heterogenous markups in a neoclassical growth model and an endogenous growth model, respectively. Edmond et al. (2015) and Arkolakis et al. (2019) analyze this effect in the international trade context. This paper extends the discussion into the patent market in an endogenous growth framework. The novel part is that the endogenous response of inventors, shifting their research directions to make patents more appealing to monopolists, amplifies the distortion from the patent trade significantly.

The paper is organized as follows. In section 2, after introducing the institutional background of the 2008 tax reform, we document several salient empirical facts about the patent trade. Section 3 builds the model under the inspiration of these motivation facts. In section 4, we apply the model to the Chinese micro-data and quantify the welfare impact of the tax reform. Finally, section 5 concludes.

2 Salient empirical patterns

In this section, we present the 2008 tax reform in China, which resulted in a significant reduction of taxes in the patent trade. Using this tax reform as an exogenous shock, we highlight four salient empirical facts regarding the patent trade.

2.1 Background and data

In China, the sale of patents is subject to income tax, with varying rates based on the seller's classification as a corporate firm or individual. Since the 1990s, personal income tax in China has been set at 25%, whereas prior to 2007, corporate income tax followed a dual-track system, with domestically-owned firms paying a base rate of 33% and foreign-owned firms paying rates ranging from 15% to 24%. In 2008, a corporate income tax reform was implemented to abolish the dual-track system. Since then, the income tax rate for both domestic-owned and foreign-owned firms has been set to 25%. As a result, domestically-owned corporate patent holders experience greater tax benefits when selling patents.

In this paper, we use two data sources: patent assignment data and Chinese manufacturing

firm survey data from 2004 to 2011. The first dataset tracks the ownership of patents and any transfers (assignments) made through trade. For each patent, we compile information such as the grant year, industry classification, cited and citing patents, and, in case of a trade, buyer, seller, and year of the transaction. Based on their names, we determine if the buyer and seller are individuals. The second dataset is widely used. It includes entry-exit, total output, and sales information at the firm level.

In our sample, only 3% of the patents are either granted or traded by foreign-owned firms, so we exclude these transactions from our analysis. Our main empirical analysis is at the industry level. For any given industry year, we count newly granted patents and traded patents.

In the subsequent analysis, unless otherwise stated, we normalize each patent by its citation count within its cohort to adjust for quality.[10] Specifically, the weight for a patent i granted in year t is $\frac{\text{Citation}_i}{\text{Total citation}_t}$, where "Citation$_i$" is patent i's citation received three years after the patent granted, and "Total citation$_t$" is the sum of three years-forward citations for all patents granted in year t.[11]

The last two decades have seen a dramatic surge in China's patent registrations. However, less commonly recognized is the fact that the pace of patent trading has accelerated even more rapidly than the sheer volume of patents themselves. In Figure 1, the blue solid-circle line shows the patent trade count (in 10,000) from 2004 to 2011, with a notable increase in 2008. The annual growth rate of trade rose from 13% to 21% that year. As a comparison, the red dashed-square line shows the number of newly granted patents (in 10,000), with the growth rate rising from 6% to 9% after 2008. As a result, the patent trade becomes relatively larger than for newly granted patents. The ratio of traded patents to new patents increased from 7% before 2008 to 12% after 2008. At first glance, the data patterns point to 2008 as a special year. We then argue the 2008 tax reform has a significant causal impact on these patterns.

Fact 1: Patent production and trade are stimulated by the 2008 policy shock, especially in industries with many corporate patents

As explained in the introduction, our main identification argument is that since the 2008 tax reform primarily impacted corporate taxes rather than individual taxes, we can reasonably anticipate systematic variations in patent production and trade across industries, due to their distinct levels of exposure to the tax change.

[10] See Akcigit et al. (2016) for a similar adjustment.

[11] Our results are robust to different citation length choices since the three years-forward citations exhibit a strong correlation with future citations.

We first show the proportion of corporate-owned patents is a persistent and exogenous industry feature. Column 1 of Table 1 lists the average shares of corporate patents across industries. Significant variations can be observed, with the petroleum industry having 69% of its patents owned by enterprises, whereas the cultural, educational, and sporting goods industry has only 20%. Additionally, the industry-level corporate patent share displays remarkable stability, as evidenced by an annual auto-correlation coefficient of 0.96, indicating the share of corporate patents is a consistent industry characteristic. In Table 2, we provide another piece of evidence, regressing the corporate share by industry on the industry fixed effect and year fixed effect, and report the R-square value. We can see most variation of corporate-owned patents is explained by the industry fixed effect. After controlling for the industry fixed effect, the R-square value is above 0.6, while the R-square value when controlling for the year-fixed effect is much smaller.

One could speculate that industries with substantial fixed costs might exhibit higher corporate patent shares, because individuals may struggle to bear the costs of front-end experimentation. We verify this hypothesis by presenting the estimates of the economy-of-scale parameters across industries, as calculated by Lashkaripour and Lugovskyy (2018), in the second column of Table 1. The data confirms that industries with a high corporate patent share tend to have a large economy of scale. The correlation between the two columns is around 0.7.

We then use a difference-in-difference approach to estimate the impact of the tax reform on patent production and trade, taking into consideration the fact that the share of corporate patents is a pre-existing industry characteristic. The equation we estimate is as follows:

$$ln(N_{it}) = \beta_t \times \text{Corporate patent share}_{i,04} + \mu_i + \mu_t + \epsilon_{it}, \tag{1}$$

where i is an industry and t is a year. N_{it} represents the *citation-weighted* count of patents produced or traded patents in industry i in year t. "Corporate patent share$_{i,04}$" denotes the share of corporate patents in industry i in 2004, which is the first year in our sample. We control for industry and year-fixed effects, represented by μ_i and μ_t. ϵ_{it} is the error term. The coefficient β_t quantifies the impact on patent trade in year t if an industry's share of corporate patents in 2004 is 1% higher.

We plot β_t and its 95% confidence intervals in the first two graphs of Figure 2. The left figure corresponds to the patents produced, and the right graph is the patents traded. The data reveal similar information for both measures. Before 2008, no systematic difference existed in patent trade across industries, based on the share of corporate patents. However, since 2008, a 10% increase in an industry's share of corporate patents is associated with a 1% and 0.7% increase

in its patent production and trade count, respectively, compared to other industries. Because the top industry in Table 1 holds about 40% more corporate patents than the bottom industry, the estimate indicates the production and trade in the high corporate patents industry are about 4% and 2.8% greater, respectively, than the low corporate patents industry. This difference is substantial, considering that post-2008 patent production and trade increased by 3% and 8% more, respectively, than the pre-2008 trend.

Because the tax reform directly targets firms, we expect patents produced or sold by firms to exhibit a more pronounced pattern than individual inventors within the same sector. We then modify the left-hand-side variable of equation (1) to patents produced or sold by firms, and plot the corresponding β_t in the second row of Figure 2. The data pattern confirms our conjecture: industries with higher shares of corporate patents exhibit a larger increase in patent production and trade only since 2008. Moreover, the magnitude of this increase is more than 2 times bigger than in the left graph, suggesting the rise in patent production and trade is largely driven by firms as producers or sellers, who benefit more from tax reform than individual inventors.

Fact 2: New patents show closer proximity to big firms, especially in industries with many corporate patents

As pointed out by Akcigit et al. (2016), firms are more likely to acquire patents that are closely connected to their current patents. Inspired by this finding, for each newly granted patent, we are interested in its relationship to the incumbent firms' existing technology. We construct this measure following two steps. First, consistent with the existing literature (Akcigit et al. (2016) or Bloom et al. (2002)), we use the patent citation network to quantify the connection between a new patent i granted in year t and the other existing patent j. The citation vector for patent i, X_i, is a representation of the shares of patents it cites in each technology category (2-digit international patent classification (IPC) code), with the constraint that $\sum_{IPC} X_{i,IPC} = 1$. Similarly, we can define X_j and the connection between i and j is $1 - D(X_i, X_j)$, where $D(X_i, X_j)$ is the Euclidean distance between their citation vectors (strictly between 0 and 1).

Then, we measure the connection of the new patent i with an incumbent firm f as

$$\frac{1}{|\Omega_{ft}|} \sum_{j \in \Omega_{ft}} [1 - D(X_i, X_j)],$$

where Ω_{ft} is the set of patents owned by firm f in year t and $|\Omega_{ft}|$ is the number of patents in this set. Hence, the above equation computes the average distance of every patent j within Ω_{ft}

away from any newly granted patent i. [12]

Finally, we define the proximity of the new patent i with the incumbents' existing technology, denoted as ϕ_i, as the size-weighted average of the connection across all incumbents:

$$\phi_i = \sum_f \text{Sales share}_f \times \frac{1}{|\Omega_{ft}|} \sum_{j \in \Omega_{ft}} [1 - D(X_i, X_j)], \tag{2}$$

where Sales share_{ft} is the share of sales of firm f in the year patent i is granted. The above expression suggests that if i shares more citations with existing patents owned by a big firm, ϕ_i is greater.

Figure 3 plots the average connection of new patents with big firms, ϕ from 2004 onwards. The dashed line represents the simple equal-weighted average, while the solid line is the average of ϕ weighted by the citations a patent receives three years after being granted. Both measures show a notable increase post-2008, with the equal-weighted average rising by 18% (from 0.37 to 0.44) and the citation-weighted average increasing by 23% (from 0.3 to 0.37). This pattern reveals that newly granted patents are more likely to have similar technologies to big incumbents.

In Figure 4, we estimate an equation similar to (1) but change the dependent variable to ϕ. The left figure shows the year 2008 again is a breakpoint. Industries with many corporate patents show a stronger increase in ϕ after the shock. Since 2008, a 10% increase in an industry's share of corporate patents is associated with a 0.5% increase in ϕ relative to other industries. Moreover, this effect is stronger when we consider the patents produced by firms. This result implies inventors changed their innovation directions in response to the tax reform.

Fact 3: Big buyers and tie-connected patents trade rise in industries with many corporate patents

We then examine the patterns in the patent trade. First, we consider changes in the buyers' characteristics. Our focus is on determining whether the tax reform has made patents more likely to be acquired by monopolists (i.e, high-market-share firms within an industry). We then estimate an equation similar to (1), but with the dependent variable being the *citation-weighted* count of patents purchased by firms whose sales are within the top 10% for a given industry and year. The time-series plot of β_t and its 95% confidence intervals are displayed in the left graph of Figure 5.

[12] To link the patent owner information from the patent assignment data with the firm sales information from the manufacturing firm survey data, we use firm names. To ensure accurate linking, we clean and standardize the firm names by performing the following steps: (1) removing all special symbols and punctuation marks that are not letters, characters, or numbers; (2) eliminating corporate forms, such as "limited corporation" or "subsidiary"; (3) converting all full-width letters and numbers into half-width ones.

The year 2008 again serves as a turning point. After 2008, in industries with 10% more corporate patents, big buyers increased by 1.1% more, whereas no such trend was observed prior to 2008. This finding implies big buyers in the industry with the largest corporate patent share increase by approximately 4.4% more than in the industry with the lowest corporate patent share.

Overall, the share of patent trade with buyers having sales in the top 10th percentile is 19% before 2008 and 25% after 2008. Compared with this number, the change in big buyers across industries because of the tax reform is substantial.

Because firms prefer to acquire closely connected patents (Akcigit et al. (2016)), we conjecture that big buyers tend to buy patents with higher ϕ. To test this hypothesis, we estimate equation (1), with the dependent variable being the count of traded patents with a connection measure ϕ greater than the median and purchased by the top 10% firms. The right graph of Figure 5 shows the plotted β_t. Compared with the left graph, it shows the trade involving big buyers and tie-connected patents rises more significantly. Industries with 10% more corporate patents show an extra increase of about 1% after 2008.

As a robustness check, we also use whether the firm's markup, defined as firm accounting revenue over total costs, is above the top 10th percentile as a measure of the buyer's monopoly power. Our results remain consistent with this alternative measure.

Fact 4: Patent trade raises markup more than productivity

We then delve into the examination of the impact of increased patent trade on measures of aggregate welfare. As noted in prior research, the impact on welfare can stem from two main channels: changes in productivity and monopoly power caused by trade.[13] We begin by analyzing some summary statistics regarding the changes that took place before and after the year 2008. In the pre-2008 period, 58% and 61% of transactions respectively demonstrated an increase in TFP and markup.[14] However, after 2008, a small increase in the proportion of transactions resulted in increased TFP (59%) and an increase in the proportion of transactions, leading to improved markup (69%).

To establish a causal connection between patent trade and its impact, we use the following empirical strategy:

$$Y_{it} = \beta ln(Trade_{it-1}) + \mu_i + \mu_t + \epsilon_{it}, \tag{3}$$

[13]For instance, Akcigit et al. (2016) focuses on the productivity gain in the patent trade, and Serrano (2010) focuses on the monopoly power change in the patent trade.

[14]We measure the industry real TFP by the method proposed by Olley and Pakes (1996). using the industry price index as the deflator, and the industry markup by the total revenue over the total cost.

where i is an industry and t stands for a year. Y_{it} is the average TFP, average markup, or the Herfindahl index in industry i in year t. $Trade_{it}$ is the *citation-weighted* number of patents traded in industry i, and β denotes its effect on Y_{it}. The model also takes into account industry-specific and year-specific effects, represented by μ_i and μ_t, respectively. The error term, ϵ_{it}, represents any unobserved factors that may affect the outcome.

To address the endogeneity problem in the above specification, we instrument $Trade_{it}$ with $IVTrade_{it}$ as:

$$IVTrade_{it} = \frac{\text{Patents owned by firms}_{i,04}}{\text{Total patents}_{i,04}} \times \text{Taxcut}_t, \tag{4}$$

where $\frac{\text{Patents owned by enterprises}_{i,04}}{\text{Total patents}_{i,04}}$ represents the proportion of patents owned by corporations in the year 2004. Taxcut_t is the tax cut in year t. It equals 0 before 2008, and 8% (33%-25%) since 2008. The IV utilizes the observation that those industries with many corporate patents receive greater trade tax reductions in the reform. Given that the initial distribution of patents among firms and individuals across industries is exogenous, we can identify the causal effects of trade on Y_{it} by contrasting industries with various exposures to patent trade tax reductions prior to and after 2008.

In the first-stage estimation, we find the coefficient $\alpha = 0.312$ is significant at the 1% level. The F-value is approximately 28. Hence our IV strongly correlates with $Trade_{it}$.

Table 3 shows the estimation results of equation (3). Each row represents a different dependent variable, and each cell reports the estimated β in one specification. The first two columns estimate the equations (3) via OLS and IV, respectively. From column 2, a 10% increase in last year's trade can improve the industry average TFP by 1.4% but can increase the industry average markup by 2.5%. It's worth noting that in China, the average markup is about 18% and the average TFP growth rate is about 6%; these differences across industries are significant. Moreover, the patent trade raises the markup more than productivity. In a perfectly competitive market, a 1% increase in productivity would lead to a 1% rise in markup. Hence, our estimate suggests that the patent trade likely boosts the industry monopoly power, thus leading to an increase in market concentration, as demonstrated in the final row.

An important concern is that our IV is constructed as an interaction of the initial shares of corporate-owned patents across industries and the tax reduction. Our estimate of β may be biased if other shocks happening in 2008 that are correlated with pre-reform industry characteristics have a bearing on the future industry dynamics. For instance, in 2008 the Chinese government selected eight industries, which the government believes are "future industries." In these eight industries,

firms can apply for subsidies, referred to as the "InnoCom subsidies," after purchasing patents.[15] Because the InnoCom subsidy simultaneously changes the industry dynamics and the benefit from patent trade in these eight industries after 2008, it may bias our estimate. To address this bias, we exclude these eight selected industries in column 3 and find that our result changes little.

As another example, the Chinese government launched a significant stimulus policy in response to the global financial crisis, resulting in a substantial increase in funding for large state-owned enterprises (SOEs). This may have caused big SOEs to purchase more patents and alter other industry dynamics, potentially affecting industries with many SOEs or large firms more than others. To address this, we control some observed time-varying industry characteristics, including the sales share of SOEs and average sales per firm. If the 2008 stimulus policy has a significant impact, after controlling for these two variables, the estimated β would become smaller. The estimation results for each dependent variable are shown in column 4, but the results are actually the opposite. The point estimates of β become larger for TFP and average markup, while β changes little for the Herfindahl index.

The other important concern with all difference-in-difference estimates is the possibility that pre-existing trends are correlated with changes in the variable of interest. If our IV is correlated with industry-level trends in Y_{it}, the estimates may be a spurious correlation. To address this concern, we conduct a falsification test for whether changes in Y_{it} before the reform (2004-2007) are correlated with changes in predicted patent trade in the first-stage estimation after the reform (2008-2011). If our IV is correlated with pre-trends, the coefficient should be similar to those estimated with the actual pre- and post-reform data. The results, shown in column 5, indicate our findings are not driven by pre-trends, because the coefficients are insignificant in all specifications.

To summarize, we establish four key observations. These observations collectively indicate that following the tax reform, the patent trade has increased and is increasingly driven by the desire to preserve the strong monopoly power of big incumbent firms. In response to this shock, innovators are producing more patents with stronger connections to big firms, as these patents are more attractive for big firms to purchase. The subsequent questions to consider are the overall impact of a reduction in the patent trade tax on welfare and whether alternative trade subsidies could improve aggregate welfare. In the following section, we explore these questions through a quantitative model analysis.

[15]These eight industries are pharmaceutical manufacture (CSIC 27), special equipment manufacture (CSIC 36), transportation equipment manufacture (CSIC 37), communication equipment & computer manufacture (CSIC 40), precision instrument manufacture (CSIC 41), computer service (CSIC 61), software service (CSIC 62) and environmental protection industry (CSIC 80). See Wei et al. (2022) for a detailed explanation.

3 Model

3.1 Environment

The economy is endowed with L units of workers for production and L_H units of researchers. The wages of the worker and researcher are denoted by w and w_H, respectively. A measure 1 of infinitely lived representative households have preferences

$$U = \int_0^\infty e^{-\rho t} \ln (C_t)\, dt,$$

where $\rho > 0$ is the utility discount factor and C_t is the final consumption goods in time t, defined as the aggregated of a continuum of differentiated goods $i \in [0, 1]$

$$\ln C_t = \int_0^1 \ln \sum_f y_{fit} di,$$

where y_{fit} is the production of the product i by firm f. The aggregator above suggests that within a product i, products from all firms are perfect substitutes. We take C_t as the numeraire.

The budget constraint of the household is

$$C_t + \dot{B}_t = r_t B_t + w_t L + w_{H,t} L_H + \Phi_t,$$

Here B_t represents the risk-free bond and r_t is the associated interest rate. Φ_t is the lump-sum transfer received from a combination of monopoly profit and government transfer.

The production function of a firm f is given by

$$y_{fit} = z_{fit} l_{fit},$$

where z_{fit} is the productivity, and l_{fit} is the workers hiring by the firm f. For convenience, in the following analysis, we ignore footnotes f, i, and t when doing so does not cause any confusion.

We consider that each product has various generations of production technology, which is ordered in a quality ladder. For each rung on the quality ladder, the gap is $\lambda > 1$. Each generation of technology is protected by a patent and can only be used by the patent holder. Denote z^F as the frontier technology patent, and $\Delta = 1, 2, ...$ as the number of gaps between z^F and the second best patent. Δ is a sufficient statistic to describe a product.

Assume firms are competing in their prices (Bertrand competition). Then, we have that for a given product i, only the frontier the patent holder will produce as a monopolist, and the product

markup, denoted as μ, is λ^{Δ}. Then the frontier patent holder's one-period profit is

$$\pi(\Delta) = (1-\varsigma)\left(1-\frac{1}{\mu}\right)C,$$

where $\varsigma < 1$ is the firm income tax, which is collected and then lump-sum transferred back to households.

In our model, each firm can hold multiple product lines. We denote n as the number of active product lines (or best patents) a firm holds, and $[\Delta_i]_{i=1}^{n}$ as the portfolio of these products' productivity gaps between the frontier and the second-best patents. A firm can expand either through in-house innovations or by acquiring patents.

Two types of innovations Firms optimally invest in innovations. We differentiate between two types of in-house innovations: vertical and horizontal innovations. Both innovations, if adopted, can increase a product's productivity by one step on the quality ladder, but they have different impacts on the equilibrium markup.

Vertical innovation allows the current monopolist of a product to upgrade its technology by one step on the quality ladder. We assume the adoption cost for a vertical innovation by the current monopolist is zero. Therefore, once a vertical innovation is invented, the current monopolist always adopts it. In this case, the firm does not expand its product lines produced, but the productivity of the product increases from z^F to $z^F\lambda$, resulting in a markup increase from λ^{Δ} to $\lambda^{\Delta+1}$.

Conversely, horizontal innovation enables a firm to diversify its product offerings. In this scenario, the inventor is not the incumbent monopolist but rather an entity that is currently inactive in that product market. Moreover, horizontal innovation differs from vertical innovation in two other respects. First, horizontal innovation is heterogeneous in terms of its connection to the current frontier technology, represented by $\phi \in [0,1]$, and inventors must select ϕ optimally.[16] Second, horizontal innovation is not free to adopt. The inventor, having no prior experience with the new product, needs to incur a one-time fixed cost to adopt the new technology.

In addition to the two aforementioned in-house innovation types, firms can also expand by acquiring patents from other firms. If the adoption cost of a new horizontal innovation is too large, inventors may have the incentive to sell it to other firms which can create greater value in owning the patent. Motivated by the previous empirical fact, the likelihood that a patent can be acquired increases in ϕ. We explain the patent acquisition further in the subsequent sections.

Importantly, horizontal innovation can play a role as a form of creative destruction. If a firm

[16]We can consider that all vertical innovations always have $\phi = 1$.

other than the incumbent monopolist adopts a new patent, the markup decreases from λ^{Δ} to λ, as the gap between the best patent (i.e., the newly developed invention) and the second-best patent (i.e., the original best technology) is reduced to a single step. This intensifies competition in the market. However, despite vertical innovation leading to productivity improvement of the same magnitude, the markup increases as the monopolist's advantage is magnified. The patent trade market changes aggregate productivity and the distribution of equilibrium markups by altering incentives for generating vertical and horizontal innovations and redistributing horizontal innovations among firms.

Figure 6 plots how the firm expands through either in-house innovations or external acquisitions in the model.

Note our model accommodates potential entrants who are individual inventors. Their newly developed patents are considered horizontal innovations because they have never produced any products before. Once they decide to commercialize their patents independently, they set up firms with a single active product line. Alternatively, they may opt to sell their patents and continue their work as individual inventors.

3.2 Patent trade market

In the model, only horizontal innovations will be traded.[17] The inventor faces a one-time fixed cost, $w_H d$, to commercialize the new horizontal innovation. We assume d follows a Pareto distribution with mean $\bar{d}$ and support $[d_{min}, +\infty)$. However, other firms' costs to commercialize this patent may be lower. Hence, trading the patent may be beneficial, which can correct the initial misallocation on the invention market, as emphasized by Akcigit et al. (2016) and Serrano (2010).

We consider two potential buyers for a horizontally-innovated patent: the current monopolist who is challenged by the new innovation, and the best user, who can run the new patent in the most efficient way. Both buyers compete for the inventor's patent on the trade market, and the inventor chooses the higher price offer or decides to commercialize the patent themselves.

To better explain the patent trade market, we introduce some notations. Denote $V\left([\Delta]_{i=1}^{n}\right)$ as the value of a firm with a patent portfolio $[\Delta]_{i=1}^{n}$ in period t. As it will be clear later, we pursue a solution such that $V\left([\Delta]\right) = \sum_{i=1}^{n} V\left(\Delta_i\right)$, where $V\left(\Delta\right)$ is the value of one product line with a gap Δ.

In the model, only new patents will be traded.[18] We assume the best user may arise with

[17] In other words, we assume that no other firm can run a vertical innovation patent more efficiently than its inventor (the current monopolist).

[18] In the data, over 80% of the patents traded are granted within 2 years.

probability β^X, which could be interpreted as the search friction on the trade market. Similarly, the current monopolist may meet the new patent with probability $\beta^V(\phi)$ on the trade market. The superscript "X" denotes that when the best user acquires the patent, it horizontally expands its product lines, while superscript "V" denotes the case when the current monopolist upgrades its existing product's technology vertically. We further assume that $\beta^V(\phi)$ is increasing in ϕ, which is motivated by the observation that firms are more willing to purchase patents with closer connections (Akcigit et al. (2016)).

Four scenarios are possible to occur: the new patent does not meet any buyers (with probability $(1-\beta^X)(1-\beta^V(\phi))$), the new patent only meets the best user (with probability $\beta^X(1-\beta^V(\phi))$, the new patent only meets the current monopolist (with probability $\beta^V(\phi)(1-\beta^X)$), and the new patent meets two buyers simultaneously (with probability $\beta^V(\phi)\beta^X$).

If the inventor is unable to sell the patent, as in the first scenario, it gets $[V(1)-w_H d]^+$, where $[.]^+$ denotes the inventor's optimal commercialization choice. We define σ, the probability that the inventor can commercialize the patent, as $\sigma = Pr(w_H d < V(1))$.

Acquisition technology Without loss of generality, we normalize the best user's fixed cost to commercialize the patent to be 0. Hence if the best user acquires the patent, it will always adopt the new technology, resulting in the frontier productivity moving from z^F to $z^F\lambda$, while the markup changes from λ^Δ to λ.

However, the current monopolist may not be able to run the new patent efficiently. Upon encountering the seller, the current monopolist draws a random adoption cost $w_H d^V \geq 0$. Generally, we assume the distribution of d^V, $f(d^V|\phi,\Delta)$, depends on ϕ and Δ, with the support $[0,+\infty)$. When $d^V = 0$, the monopolist can run the new patent as efficiently as the best user. Note that the current monopolist may not want to adopt the new invention if d^V is too high, but it still has the incentive to buy the patent, otherwise, it will have lost the leading position.

Patent price After observing d^V and d, the buyers offer prices. Note the inventor is willing to sell the patent only if price $p \geq \frac{[V(1)-w_H d]^+}{1-\gamma}$, where $[V(1)-w_H d]^+$ is the value when the inventor keeps the patent. γ is the tax rate when selling the patent. In the Chinese context, $\gamma = \varsigma$ if the seller is a firm; otherwise, $\gamma = \varsigma_{\text{individual}}$, the individual income tax.

If only one buyer is present, the buyer will just offer a price $p = \frac{[V(1)-w_H d]^+}{1-\gamma}$ to make the seller indifferent between keeping the patent and selling it. In this case, the inventor, after paying the tax, always gets $[V(1)-w_H d]^+$, regardless of who is the buyer.

We further assume the trade tax never deters the best user from acquiring the patent: $V(1) \geq \frac{[V(1)-w_H d_{min}]^+}{1-max\{\varsigma,\varsigma_{\text{individual}}\}}$. In other words, when the best user is the only potential buyer, under any possible

realization of d, it is always optimal for the best user to purchase the patent, regardless of whether the seller is firm. The best user gets

$$V(1) - \frac{[V(1) - w_H d]^+}{1-\gamma}.$$

If the current monopolist is the only buyer, it gets

$$\max[V(\Delta+1) - w_H d^V, V(\Delta)] - \frac{[V(1) - w_H d]^+}{1-\gamma}.$$

Here, the max operator takes into account the optimal technology adoption decision: if d^V is too large, the monopolist chooses to bury the new patent. As $\Delta \geq 1$, the above equation is always positive. The current monopolist would always want to purchase the patent upon meeting the new patent, disregarding the realization of the d^V.

If two buyers simultaneously show up, the price equals the limit price as a Betrand competition:

$$p = \min[\underbrace{\max[V(\Delta+1) - w_H d^V, V(\Delta))]}_{\text{optimal tech. adoption of the monopolist}}, V(1)].$$

Since $\Delta \geq 1$, the monopolist can always offer a higher price than the best user for any d^V. Hence, in this case, the monopolist will always get the patent under the price $p = V(1)$.

Expected gain from trade market We then define the expected gain from trade for the two buyers and the seller. For the seller with a new horizontal innovation patent ϕ, denote its expected value from entering the trade market as $\Omega(\phi; \varsigma)$:

$$\Omega(\phi;\gamma) = (1 - \beta^V(\phi)\beta^X)E_d(V(1) - w_H d)^+ + (1-\gamma)\beta^V(\phi)\beta^X V(1). \tag{5}$$

If only one buyer shows up, the seller always receives $E_d[V(1) - w_H d]^+$ after considering the trade tax. When two buyers appear simultaneously, the trade will always happen, and the pre-tax price is $V(1)$.

We see the value of a horizontal innovation $\Omega(\phi;\gamma)$ increasing in the patent connection ϕ since the patent price is higher, but decreasing in the income tax ς. In particular, the patent trade tax only changes $\Omega(\phi;\gamma)$ through the second term: the scenario with two buyers. When the patent trade tax is lower, the value of a large ϕ patent increases more.

We then switch to the current monopolist's gain from the trade. Let $B^V(\Delta)$ represent the current monopolist's expected value from a "monopoly-power-driven" acquisition if a new patent

challenges the firm's leading position:

$$B^V(\Delta) = E_{\phi,\gamma,d^V,d}\{\beta^V(\phi)\underbrace{[(1-\beta^X)[\max(V(\Delta+1)-w_Hd^V,V(\Delta))-\frac{[V(1)-w_Hd]^+}{1-\gamma}]}_{\text{the current monopolist as the single buyer}}+ \underbrace{\beta^X[\max(V(\Delta+1)-w_Hd^V,V(\Delta))-V(1)]]}_{\text{2 buyers}}\}, \tag{6}$$

where the expectation, $E_{\phi,\gamma,d^V,d}$, is taken over the equilibrium distribution of ϕ and all possible realizations of d^V and d after considering the possibility of the seller being firm or not. The first and second lines represent the cases where only the current monopolist wants to buy the patent and when two buyers meet the seller simultaneously.

Because the current monopolist can always acquire the patent if it meets the seller, the expected chance of a "monopoly-power-driven" acquisition, denoted as θ^V, is

$$\theta^V = E_\phi[\beta^V(\phi)]. \tag{7}$$

Here the expectation is taken over the equilibrium distribution of ϕ. We can also have the probability that the current monopolist is willing to adopt the newly acquired patent:

$$\sigma^V(\Delta) = E_{d^V}[Pr(w_Hd^V < V(\Delta+1)-V(\Delta))|\Delta], \tag{8}$$

where the expectation is taken over the distribution of the adoption cost d^V conditional on Δ. If d^V increases with Δ, the monopolist with a larger Δ is more likely to suppress the newly acquired horizontal technology. This not only helps the monopolist maintain a high markup but also harms productivity.

For the best user, its expected value from a "productivity-driven" acquisition, B^X, is

$$B^X = E_{\phi,\gamma,d}\beta^X\underbrace{[(1-\beta^V(\phi))(V(1)-\frac{[V(1)-w_Hd]^+}{1-\gamma})]}_{\text{the best user as the single buyer}}. \tag{9}$$

The best user will always adopt the newly acquired patent and the chance it can get the patent, θ^X, is

$$\theta^X = E_\phi[\beta^X(1-\beta^V(\phi))] \tag{10}$$

where we use the fact that the best user can only get the patent if the current monopolist does not show up.

If new horizontal innovations have a higher ϕ, then we observe an increase in the frequency

of "monopoly-power-driven" trade (θ^V), as well as an increase in the likelihood that current monopolists adopt the technology ($\sigma^V(\Delta)$). At the same time, the frequency of "productivity-driven" trade (θ^X) declines.

3.3 Firm problem and equilibrium

We focus on the balanced growth path (BGP), on which the growth rate is g, and the aggregate horizontal innovation per period is τ. The household optimality condition yields that the interest rate $r = \rho + g$.

On the BGP, the product value could be characterized by the following Bellman equation:

$$\begin{aligned}(\rho+g)V(\Delta) \;=\; & \max_{I(\Delta),x(\Delta)} \underbrace{(1-\varsigma)\pi(\Delta)}_{\text{1-period profit}} + \underbrace{I(\Delta)[V(\Delta+1)-V(\Delta)] - w_H C^V(I(\Delta))}_{\text{internal vertical innov.}} \\ & + \underbrace{[x(\Delta)\Gamma - w_H C^X(x(\Delta))]}_{\text{internal horizontal innov.}} + \\ & \underbrace{\tau[B^V(\Delta)-V(\Delta)]}_{\text{"monopoly-power-driven" acquisition}} + \underbrace{\tau B^X}_{\text{"productivity-driven" acquisition}}\end{aligned} \tag{11}$$

Here $I(\Delta)$ and $x(\Delta)$ denote the optimal arrival rates of internal vertical and horizontal innovations for a given product, and $C^V(I)$ and $C^X(x)$ are the associated research costs (measured by the number of researchers). Γ is the expected value of a successful horizontal innovation, which we define later. The firm optimally chooses $I(\Delta)$ and $x(\Delta)$. The last two terms are the values from "monopoly-power-driven" and "productivity-driven" acquisitions.

Note that $I(\Delta)$ and $x(\Delta)$ are defined at the product level. Since a firm has n products, the firm-level innovation needs to aggregate x and I across all products within the firm.

When considering the horizontal innovation, the firm first randomly draws a menu of cost $\epsilon(\phi)$ for each possible ϕ, and then chooses an optimal ϕ to maximize Γ:

$$\Gamma = E_{\epsilon(\phi)} \max_{\phi} \left[\Omega\left(\phi;\varsigma\right) - w_H \epsilon(\phi)\right]. \tag{12}$$

Here the expectation is taken over the random innovation cost $\epsilon(\phi)$, which follows a Type-I Gumbel distribution. Denote $\omega(\phi)$ as the share of firms' horizontal innovations with connection ϕ,

$$\omega(\phi) = \frac{exp(\Omega(\phi;\varsigma)/w_H)}{\sum_{\phi'} exp(\Omega(\phi';\varsigma)/w_H)} \tag{13}$$

$$\Gamma = w_H \ln[\sum_{\phi} exp(\Omega(\phi;\varsigma)/w_H)]. \tag{14}$$

An inventor will keep its invented patent with probability $\eta = \sum_{\phi}[\omega(\phi)(1-\beta^X)(1-\beta^V(\phi))]$.

For the individual innovator, the free-entry condition requires that

$$E_{\epsilon(\phi)} \max_{\phi} \left[\Omega\left(\phi; \varsigma_{\text{individual}}\right) - w_H \epsilon(\phi)\right] = w_H v^E, \tag{15}$$

where v^E is the entry cost and $\varsigma_{\text{individual}}$ is the individual income tax rate. Similarly, the share of individuals' horizontal innovations with connection ϕ, $\omega_{\text{individual}}(\phi)$, and v_E satisfy

$$\omega_{\text{individual}}(\phi) = \frac{exp(\Omega(\phi; \varsigma_{\text{individual}})/w_H)}{\sum_{\phi'} exp(\Omega(\phi'; \varsigma_{\text{individual}})/w_H)} \tag{16}$$

$$v^E = \ln[\sum_{\phi} exp(\Omega(\phi; \varsigma_{\text{individual}})/w_H)]. \tag{17}$$

Due to the tax rate difference, the probability that an individual inventor will keep her invented patent is $\eta_{\text{individual}} = \sum_{\phi}[\omega_{\text{individual}}(\phi)(1-\beta^X)(1-\beta^V(\phi))]$.

In the equilibrium, the aggregate horizontal innovation τ comes from both individual innovators and the firm's horizontal innovation. The entry rate of new inventors is denoted as u, and

$$\tau = u + \int_0^1 x_i di, \tag{18}$$

where x_i is the horizontal innovation intensity of the monopolist firm for product i. Note that only some of these horizontal patents will be adopted. If the patents are not sold, σ share of them will be adopted. If the patents are sold to the monopolists, σ^V share will be adopted. But if the patents are sold to the best user, all of them will be adopted.

On the patent trade market, we have the total patents sold equal to the total patents acquired:

$$u(1-\eta_{\text{individual}}) + \int_0^1 x_i di(1-\eta) = \tau\theta^X + \tau\theta^V. \tag{19}$$

3.4 Characterize the equilibrium

To characterize the equilibrium, we impose a few functional form assumptions on the innovation cost functions: $C^V(I) = \lambda^{-\Delta}\frac{c^V}{1+\zeta}I^{1+\zeta}$ and $C^X(x) = \frac{c^X}{1+\zeta}x^{1+\zeta}$, where c^V and c^X are the cost shifters of the two types innovations. For $C^V(I)$, the functional assumption suggests that if the monopolist has a large advantage Δ, its innovation cost becomes lower. This assumption helps keep the problem tractable and is widely used in the literature.

We then assume that the adoption cost of the current monopolist has a functional form of

$$d^V = \lambda^{-\Delta}\alpha(\phi)\tilde{d^v}, \tag{20}$$

where $\tilde{d^v}$ follows an exponential distribution with the mean normalized as 1. The above equation suggests that the adoption cost declines with Δ with a rate λ. $\alpha(\phi)$ represents the component in the adoption cost that correlates with connection ϕ. We assume $\alpha(\phi)$ is decreasing in ϕ so that a tie-connected patent is less costly to adopt.

We first characterize the solution of the firm problem (11). The appendix shows the following result.

Proposition 1. (1) Under the BGP, the firms' horizontal innovation intensity x and the probabilities to acquire patents θ^V and θ^X are three constant and do not depend on Δ.

(2) The vertical innovation intensity I and adoption probability of acquired patents σ^V are constants and do not depend on Δ.

Thus, the value function has a similar property to many other creative-destruction models (Klette and Kortum (2004)): the firm expansion rate follows Gibrat's law.

Next, we characterize how the tax reform can change the horizontal innovation incentive. From equation (5), the trade tax will only change the payoff of the seller in the case of two buyers simultaneously arising. Since a higher ϕ increases the chance of this scenario, we can have the following result.

Proposition 2. A reduction in the firm income tax ς induces $\omega(\phi)$ and $\omega_{\text{individual}}(\phi)$ increases (first-order stochastic dominance), while the increase in $\omega(\phi)$ are more pronounced.

In other words, a reduction in income tax can ease the difficulty of patents sold to current monopolists, both for individual-owned patents and firm-owned patents. Since firm inventors receive more direct effects from the tax reduction, their changes in innovation investments are more significant and their patents are more likely to sell, which is consistent with the empirical patterns documented before.

3.4.1 Aggregate welfare effect of patent trade

We are interested in how a reduction in trade tax changes the aggregate welfare of the economy. In this section, we first introduce the aggregate welfare measure of the economy and then show how the trade tax reduction can change various sources of aggregate welfare.

From the labor market clearing condition, we can have that the labor share of the economy is determined by the average inverse of markup:

$$\frac{w_t L}{Y_t} = \sum_{\Delta} \left(\frac{1}{\mu(\Delta)} \right) h(\Delta). \tag{21}$$

Here $h(\Delta)$ is the equilibrium distribution of product technology gap Δ and $\mu(\Delta) = \lambda^{\Delta}$ is the product markup. Thus when firms have stronger monopoly power, the labor share is lower as monopoly profit increases.

Then the aggregate production function can be written as

$$Y_t = \underbrace{Z_t}_{\text{technology frontier}} \times \underbrace{M}_{\text{Misallocation due to monopoly power}} \times L, \tag{22}$$

where

$$\ln Z_t = \int \ln z_{it}^F di \tag{23}$$

$$M = \frac{\exp \sum_{\Delta} \left(\ln \frac{1}{\mu(\Delta)} \right) h(\Delta)}{\sum_{\Delta} \left(\frac{1}{\mu(\Delta)} \right) h(\Delta)}. \tag{24}$$

The aggregate TFP of the economy, $\frac{Y_t}{L}$, is determined by two components. The first term is the level of the aggregate technology frontier, $\ln Z_t$, which increases with a rate g on the BGP. The second component, M, captures the misallocation due to the monopoly power. We can easily show that $M \leq 1$, and that $M = 1$ if and only if markups are equalized. This result is highlighted in the misallocation literature (Edmond et al. (2018)): the dispersion of markup creates a loss in the aggregate TFP.

The aggregate welfare in the economy then is defined as

$$\text{Welfare} = \int_0^{\infty} e^{-\rho t} \ln C_t dt \propto \frac{1}{\rho} \ln M + \frac{1}{\rho - g}. \tag{25}$$

We then characterize g and M, separately. Firstly, the aggregate growth rate on the BGP follows

$$g = ln(\lambda) \times [\underbrace{\tau\sigma(1 - \theta^X - \theta^V) + \tau\theta^X}_{\text{growth with lower markup}} + \underbrace{I + \tau\theta^V\sigma^V}_{\text{growth with higher markup}}]. \tag{26}$$

Equation (26) highlights that the growth of the economy arises from four sources. First, aggregate non-traded horizontal innovation is $\tau(1-\theta^X-\theta^V)$, among which the σ share is adopted and raises the productivity. Second, the aggregate measure of horizontal patents sold to best users is $\tau\theta^X$. All of them will be adopted. In these two scenarios, creative destruction results in

current monopolists losing their leading positions. We denote $\hat{\tau}$ as the creative-destruction rate, the rate at which the current monopolist will be replaced by new firms:

$$\hat{\tau} = \tau\sigma(1 - \theta^X - \theta^V) + \tau\theta^X. \tag{27}$$

The third term of equation (26) represents the aggregate internal vertical innovation. In the last term, $\tau\theta^V$ is the aggregate measure of patents sold to current monopolists, among which only σ^V will be adopted. In these two cases, the growth comes at the cost of a higher markup. The sum of these two terms is $I + \tau\theta^V\sigma^V$.

From equation (26), we can see the trade-off when encouraging patent trade. Encouraging "productivity-driven" trade is always wise because an increase in θ^X increases aggregate growth, due to correcting initial misallocation on the invention market ($\sigma < 1$), but without any distortion on markup. However, if the "monopoly-power-driven" trade increases, the welfare effect is not clear. It depends on the trade-off between the change in productivity and the distortion in markup. In the worst case, if $\sigma^V < \sigma$, the "monopoly-power-driven" trade not only increases the markup, but also hurts the aggregate growth since monopolists are less likely to adopt the new technology.

We can also characterize $h(\Delta)$, the stationary distribution of Δ at the product level. Note that on the BGP, $h(\Delta)$ satisfies

$$(I + \tau\theta^V\sigma^V + \hat{\tau})h(\Delta) = (I + \tau\theta^V\sigma^V)h(\Delta - 1), \text{ if } \Delta \geq 2 \tag{28}$$

$$(I + \tau\theta^V\sigma^V + \hat{\tau})h(\Delta) = \hat{\tau}, \text{ if } \Delta = 1.$$

The left-hand side in equation (28) is the measure of products whose productivity gaps do not keep at level Δ, due to internal vertical innovation, external acquisition, or creative destruction. The right-hand side is the measure of products whose productivity gaps increase from $\Delta - 1$ to Δ. The appendix shows the following result.

Proposition 3. The stationary product-level markup distribution follows a Pareto distribution with a shape parameter χ.

$$Pr(\mu > \bar{\mu}) \propto \bar{\mu}^{-\chi} \text{ ,where } \chi = \frac{1}{ln\lambda} ln(1 + \frac{\hat{\tau}}{I + \tau\theta^V\sigma^V}). \tag{29}$$

The above proposition states that the dispersion of Δ increases with the ratio $\frac{\hat{\tau}}{I+\tau\theta^V\sigma^V}$. If the "monopoly-power-driven" trade is more frequent (an increase in θ^V) or the monopolist is more likely to adopt the newly acquired patent (an increase in σ^V), the current monopolist is more likely to acquire patents to maintain its leading position. Hence the markup distribution gets a

fatter tail, which will result in a higher misallocation in the aggregate welfare. Meanwhile, when θ^X increases or τ/I increases, the tail is thinner.

From Proposition 2, we can see a reduction in income tax changes the aggregate welfare from three channels. First, the patent trade can change the incentives of new innovations, bringing more close-connected patents with monopolists (increase in τ). Second, due to the increase in ϕ, patents are easier to be sold to monopolists. Hence, θ^V increases but θ^X declines, which will cause M to decline due to the increase in the dispersion of markup (proposition 3). Meanwhile, the average adoption rate of acquired patents will decline so that the productivity gain from trade is smaller. Third, the general equilibrium effect will also impact in-house vertical innovation. On the one hand, an increase in horizontal innovation may attract more researchers to divert to inventing or adopting horizontal innovations, potentially leading to an increase in w_H. On the other hand, trade changes the rate of creative destruction, which will affect vertical innovation investment.

3.4.2 Cross-sectional distribution of firm size, age, and markup

We then explore the impacts of patent trade on the cross-sectional distribution of firm size, age, and markup. Note a firm is a collection of products, whose firm-level markup, μ_f, is the average markup across its products:

$$\mu_f = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \lambda^{-\Delta_i}}.$$

Meanwhile, due to the log aggregator assumption in the household problem, the sale of each product is the same. Hence a firm's total sale is proportional to its number of products. The firm's sales growth is determined by the probability of adding and losing a product.

Note that the probability of adding a product is $x\eta\sigma + \tau\theta^X$. The first term represents the horizontal innovations that are kept and commercialized by the firm itself. The second term represents the acquired patents as "best users." Meanwhile, the probability of losing a product is simply $\hat{\tau}$. Hence the ratio between these two rates determines the distribution of the number of products at the firm level. As Klette and Kortum (2004) show, we can have the following result:

Proposition 4. (Klette and Kortum (2004)): The ratio between the sale of firm f and the average firm sale Y, $\frac{Sales_f}{Y}$ has a distribution

$$\lim_{s\to\infty} Pr(\frac{Sales_f}{Y} \geq s) \propto \frac{1}{s}(\frac{x\eta\sigma + \tau\theta^X}{\hat{\tau}})^s.$$

The tail of the firm size distribution is fat, converging much slower than an exponential distribution or normal distribution. After a reduction in firm income tax, two forces work in opposite directions. On one hand, firms invest more in horizontal innovation and face more patents for

sale (increase in $\frac{x}{\hat{\tau}}$ and $\frac{\tau}{\hat{\tau}}$), and firm size tends to be larger. On the other hand, the increase of "monopoly-power-driven" trade decreases η and θ^X and the firms are less likely to grow. In the end, these two forces determine whether the market concentration would increase.

We can further characterize the firm sales and firm markup conditional on firm age. Following the insight of Peters (2020), the firm size and markup conditional on firm age satisfy the following.

Proposition 5. (Peters (2020)): Conditional on firm age a_f, (1) the firm survival rate is

$$\text{Survival}_f(a_f) = 1 - \frac{\hat{\tau}}{x\eta\sigma + \tau\theta^X}\gamma(a_f),$$

where $\gamma(a) = \frac{(x\eta\sigma+\tau\theta^X)(1-e^{-(\hat{\tau}-x\eta\sigma-\tau\theta^X)a})}{\hat{\tau}-(x\eta\sigma+\tau\theta^X)e^{-(\hat{\tau}-x\eta\sigma-\tau\theta^X)a}}$.

(2) The firm sales conditional on a_f follow a distribution

$$Pr(\frac{\text{Sales}_f}{Y} > s|a_f) = \gamma(a_f)^s(1-\gamma(a_f)).$$

(3) The firm log markup conditional on a_f is proportional to the average productivity gap across all products the firm has:

$$E(\ln \mu_f|a_f) \propto \sum_{\Delta=1}^{\infty}\int_0^{a_f} \Delta\frac{m(\Delta|a_P)}{1-m(0|a_P)}g(a_P|a_f)da_P.$$

Here $m(\Delta|a_P)$ is the distribution of Δ conditional on product age a_P, which follows

$$m(\Delta|a_P) = \begin{cases} \frac{1}{(\Delta-1)!}(I+\tau\theta^V\sigma^V)^{\Delta-1}a_P^{\Delta-1}e^{-(I+\tau\theta^V\sigma^V+\hat{\tau})a_P} & \text{if } \Delta \geq 1 \\ 1-e^{-\hat{\tau}a_P} & \text{if } \Delta = 0. \end{cases}$$

$\frac{m(\Delta|a_P)}{1-m(0|a_P)}$ is the distribution of Δ as a function of product age a_P conditional on the product surviving. $g(a_P|a_f)$ is the distribution of product age a_P conditional on firm age a_f,

$$g(a_P = a|a_f) = \begin{cases} g_0(a_f) & \text{if } a_P = a_f \\ (1-g_0(a_f))g_1(a_P|a_f) & \text{if } a_P < a_f, \end{cases}$$

where $g_0(a_f) = e^{-(x\eta\sigma+\tau\theta^X+\hat{\tau})a_f}\frac{1}{\gamma(a_f)}ln(\frac{1}{1-\gamma(a_f)})$ is the probability that a firm with age a_f has the initial product ($a_P = a_f$). $g_1(a_P|a_f) = \frac{\hat{\tau}e^{-\hat{\tau}a_P}+(x\eta\sigma+\tau\theta^X)e^{-(x\eta\sigma+\tau\theta^X+\hat{\tau})a_f}e^{(x\eta\sigma+\tau\theta^X)a_P}}{1-e^{-(x\eta\sigma+\tau\theta^X+\hat{\tau})a_f}}$ is the distribution product age conditional on it not being the initial product.

From the above proposition, two ratios jump out. First, the ratio between adding a product and losing a product, $\frac{x\eta\sigma+\tau\theta^X}{\hat{\tau}}$, determines the distributions of firm sales, the survival rate, and

the product age.

Second, the firm markup is the average of all products' markups. For a given product, its markup is determined by the ratio that the monopolist can increase Δ: $I + \tau\theta^V\sigma^V$. Hence the distribution of firm-level markup is jointly determined by these two ratios. When θ^V increases and θ^X declines, firms tend to have older products and higher markups.

3.4.3 Changes in firm-level markup and sales after acquiring patents

Next, we characterize the expected changes in firm markup and sales after acquiring a patent, as

$$d\ln\mu_f \propto \frac{\theta^X}{\theta^V+\theta^X}\frac{1}{\text{sales}_f}(1-\lambda^{-1}\mu_f) + \frac{\theta^V}{\theta^V+\theta^X}\frac{1}{\text{sales}_f}(\lambda^{-1}\sigma^V)$$

$$d\ln\text{sales}_f \propto \frac{\theta^X}{\theta^V+\theta^X}\frac{1}{\text{sales}_f}.$$

The first equation has two components. The first term is the change in $\ln\mu_f$ in a "productivity-driven" acquisition. Note $1-\lambda^{-1}\mu_f < 0$. So when μ_f is higher, a "productivity-driven" acquisition cuts firm markup more, because the newly acquired product has a much smaller markup than the existing products. The second term adjusts the firm markup for a "monopoly-power-driven" acquisition. $\frac{\theta^V}{\theta^V+\theta^X}$ is the ratio of a "monopoly-power-driven" acquisition, while $\lambda^{-1}\sigma^V$ is the expected increase of $\sum_{i=1}^n \lambda^{-\Delta_i}$ from this acquisition. This equation indicates that when "monopoly-power-driven" acquisitions are more frequent, the increase in firm markup is greater. Meanwhile, the increase is more pronounced for high markup firms, since the patent trade now is more likely to make "stronger firms even stronger."

Because the "productivity-driven" acquisition increases firm sales and tends to decline the firm's markup, the second equation indicates that, given the firm sales before an acquisition, if the "productivity-driven" acquisitions are less frequent, the increase in firm sales after acquiring a patent is smaller.

Moreover, the above two equations indicate that the impact of an acquisition on the firm is decreasing with firm size since each acquisition only changes one product line.

When the firm income tax rate ς decreases, the likelihood of "monopoly-power-driven" trade increases. Therefore, our model predicts that the increase in buyer markup will be more pronounced after tax reform, as shown in the empirical facts. This effect is particularly strong for firms with high markups before the acquisition. We use this prediction to validate our model later.

4 Quantitative Analysis

4.1 Calibration

We use the estimates in the literature to guide the choice of the parameter values. For those not available in the literature, we calibrate their values to match some key moments in the data. We only use the data before the tax reform to compute targeted moments. We then evaluate the welfare of the policy by decreasing the tax rate in the same way as the 2008 tax reform.

Pre-assigned parameters

We set the domestic owned firms' income tax rate, ς, at 33%, and the individual income tax, $\varsigma_{\text{individual}}$, at 25%. We also set the discount factor $\rho = 0.05$ and $L = 1$, and calibrate H to match the observed relative salary between college graduates and other workers. Following Acemoglu and Akcigit (2012), we pick the curvature of the innovation cost function $\zeta = 1$.

Next, the minimal adoption cost for inventors, d_{min}, is assumed to make the best user get 0 when acquiring patents ($V(1) = \frac{[V(1)-w_H d_{min}]^+}{1-\varsigma}$). In this case, the best user will always want to acquire patents for any possible realization of d.[19] We then assume the adoption cost for monopolist, d^V, declines in ϕ (as shown in equation (20)), $\alpha(\phi) = \alpha_0(1-\phi)^{-\alpha_1}$. Hence when $\phi = 1$, $d^V = 0$. If $\alpha_1 < 0$, the adoption cost is lower for high ϕ patents.

Finally, we assume the probability of meeting the current monopolist $\beta^V(\phi) = \beta_0\phi^{\beta_1}$. If $\beta_1 > 0$, the chance to sell the patent is higher when ϕ is higher. We want to emphasize that when we calibrate the model, we do not impose α_1 and β_1 to be positive, but allow the data to discipline the parameters directly.

Other parameters and identification

We calibrate 11 parameters: (1) step size λ; (2) two innovation cost parameters c^V and c^X; (3) entry cost for new entrants v^E and cost of commercializing its horizontal innovation $\bar{d}$; (4) five parameters associated with patent acquisition: β_0, β_1, β^X, α_0 and α_1.

In the data, we observe firm age a_f, firm sales $sale_f$, firm markup μ_f, firm-level acquisition decision, the entry rate of new firms, the entry rate of new patents, and ϕ for each new patent. We identify each parameter as follows.

First, we choose the step size λ to match the observed aggregate TFP growth rate of the economy. Second, c^V and c^X are identified similarly to Peters (2020). c^V and c^X determine the

[19] Note that $\varsigma = 0.33$, which is higher than $\varsigma_{\text{individual}}$. So the best user will buy an individual's patent disregarding the realization of d as well.

arrival rate of vertical in-house innovation and horizontal in-house innovation. Intuitively, if c^V is smaller, vertical in-house innovation arrives more frequently. Then, the firm's markup increases in firm age in a steeper way. However, if c^X is smaller, the horizontal innovation dominates. The firm's sales increase in firm age in a steeper way. In other words, the life cycles of firm markup and sales jointly determine the frequency of two types of in-house innovations.

Third, we use the entry rate of new firms and the entry rate of newly granted patents from individual inventors to identify the entry cost v^E and the adoption cost $\bar{d}$.

The rest six acquisition technology parameters are jointly identified from trade patterns in the data. First, conditional on ϕ, the probability that a patent can be traded is

$$\Pr\left(Trade|\phi\right) = 1 - E_\phi((1-\beta^X)(1-\beta^V(\phi))).$$

Hence we calibrate β_0 and β_1 to match two moments: the trade probability for patents with the bottom 25% and top 25% observed connections, ϕ.

Second, given β_0 and β_1, β^X determine the share of "productivity-driven" trade, while α_0, and α_1 determine the gain from the "monopoly-power-driven" trade. Note that these two types of trade have different implications on a firm's markup and size. The "monopoly-power-driven" trade increases the firm's markup, while the "productivity-driven" trade decreases the firm's markup but increases firm size. Hence we calibrate them to match four moments: the change in buyer's markup and sales after acquiring patents with the bottom 25% and top 25% observed connections.

4.2 Parameter values and model fitness

We summarize all parameter values in Table 5 and report the model fitness in Table 6. Overall, the model matches the data well.

In particular, the cost of vertical innovation, c^V, is similar to the horizontal innovation cost, c^X, because the firm's markup and size have similar life cycles, so the arrival rates of the two types of innovations are similar. Compared with Peters (2020), who finds $c^V = 0.2$ and $c^X = 1.8$ in the US, our parameters suggest that the markup grows faster while firm size grows slower against firm age in China. In terms of the entry cost of a new individual innovator, we find $v^E = 0.18$, which is only about half of the value in Peters (2020). The explanation may be that we target the entry rate of new patents, whereas Peters (2020) targets the entry rate of new firms.

We then discuss the three parameters that determine the trade probability, two of which (β_0 and β_1) determine the trade probability with monopolists and the other (β^X) determines the meeting probability with the best user. $\beta_0 = 0.17$ means that when $\phi = 1$, the chance to acquire

the patent is about 17%, while $\beta^X = 0.01$ means the best user can meet the patent with 1% probability. The positive value of β_1 implies that low-connected patents are less likely to be sold.

We have four parameters associated with adoption costs. $\alpha_0 = 0.98$ means the cost of monopolists to adopt a $\phi = 0$ patent is about about two times greater than $\bar{d}$, the adoption cost of the inventor. In other words, the current monopolist is less efficient in adopting the new technology than its inventor for a low ϕ patent. Selling a low ϕ patent to the monopolist not only raises the markup but also decreases productivity. The positive value of α_1 implies that high-connected patents are more likely to be adopted by monopolists, which is consistent with the extant literature findings.

Non-targeted moments To further demonstrate the validity of our model, we explore some non-targeted moments. The first two graphs in Figure 7 depict the relationship between firm markup and sales size against firm age. The solid lines represent model-predicted relationships, and the dashed lines represent actual data. Although we focus only on the sales and markup of new entrants and five-year-old firms, our model accurately captures the overall life cycles of markups and sales. Since sales and markup are perfectly correlated for new entrants, we only target the sales growth rate.

The first graph in the second row shows the conditional survival rate of firms based on their age. Even though we do not target this moment, our model accurately predicts the firm's survival rate.

Next, the following two graphs illustrate the change in firm markup and sales relative to ϕ. The final graph displays the probability of a new patent being sold conditional on ϕ. Although we target moments at the 25th and 75th percentiles of ϕ, our model effectively captures the general patterns in the data.

4.3 Model validation

Our model predicts that when the firm income tax rate ς declines, the trade is more likely to be "monopoly-power-driven", implying that the increase in markup is more pronounced for high markup firms after controlling for firm size, as discussed in the previous section. We use this prediction as a validation test of our model.

To estimate the validation equation from the data, we use the following specification:

$$dln\mu_{f,t} = aln\mu_{f,t0} \times Post_t + ln\mu_{f,t0} + \ln sale_{f,t-1} + Post_t + X_f + \epsilon_{f,t}. \tag{30}$$

Here each observation in the estimation is a transaction. f represents the buyer and t is the acquisition year. $dln\mu_{f,t}$ is the average change in log firm markup two years before and after the transaction. $\mu_{f,t0}$ is the firm markup in the initial year (the year 2004). $Post_t = 1$ if the transaction happens after 2008 (i.e, $t \geq 2008$), and 0 otherwise. In the regression, we control for the buyer's sales before the acquisition. Other controls, X_f, include the buyer's industry code, located city, and the quality of the acquired patent, measured by the three-years-forward citation count.

The first term, an interaction between the buyer's age and $Post_t$ dummy, is our interest. Our theory predicts a positive coefficient a, suggesting that a high-markup firm enjoys a greater increase in markup in the acquisition after the tax reform.

Column 1 of Table 8 presents the estimation results. The coefficient before $\ln sale_{f,t-1}$ is significantly negative, indicating that the increase in markup is smaller for larger firms. This is consistent with our model's prediction that a single acquisition is less likely to change the markup of a large firm. The coefficient before $Post_t$ is about 0.054, implying that patent trade after 2008, on average, increases the buyer's markup.

Importantly, the coefficient before the firm's initial markup is not significantly different from 0, but the coefficient of the interaction term is significantly positive. Specifically, when a firm's initial markup increases by 1%, the increase in its markup from acquiring a patent after 2008 is 0.12% higher than the increase in markup from acquiring a patent before 2008. This finding is consistent with our model's prediction that a lower firm income tax rate increases the likelihood of "monopoly-power-driven" trade, leading to a more pronounced increase in markup, particularly for high-markup firms.

To assess how well our model captures this pattern quantitatively, we simulate our model before and after the tax reduction and estimate a regression similar to equation (30). The results, shown in column 2 of Table 8, demonstrate that the model predicts this data pattern well. In particular, high-markup firms benefit more from patent trade after the tax reform.

4.4 Assessing the tax reform

We assess the effect of China's tax reform in 2008. We reduce the corporate income tax rate, ς, from 33% to 25%, while keeping the individual income tax rate and all other parameters unchanged. We then re-compute the balanced growth path. In panel A of Table 7, columns 1 and 2 compare some key moments before and after the tax reform. As expected, although both in-house vertical and horizontal innovation increase, the latter does so more significantly, because the tax cut provides an extra incentive for firms to generate and sell patents but not for individual

inventors.

The average ϕ of the new horizontal innovation, reported in the third row, rises by 33% $(0.39/0.29 - 1)$ following the tax cut. Patent trade has become easier, with the trade probability being doubled. However, most of the increase is "patent trade for monopoly power," with an increase in its share from 27% to 45%.

The next two rows of panel A in Table 7 compare the adoption rates of new horizontal patents and those purchased by monopolists before and after the tax reform. Before 2008, 48% of new patents were adopted, but only 13% of acquired patents were adopted when purchased by monopolists. This difference highlights the incentives of trading to maintain monopoly power rather than increase productivity. After 2008, the adoption rate of monopolists increased to 19% due to the increase in ϕ, but the overall average adoption rate declined to 42% as fewer patents were sold to the best users. Thus, the tax reduction had a negative impact on aggregate productivity since it decreased the adoption rate of newly invented patents.

The last row of panel A reports the average markup of the economy. As expected, the average markup increased by 6% (from 11% to 17%) after the tax reform.

Panel B in Table 7 decomposes the welfare of the two BGPs into several sources, as shown in equation (25). The first row corresponds to the misallocation term, which decreases due to the tax reduction. However, since more trade implies a decline in the creative destruction rate, the markup may have a fatter tail, leading to a decrease in M from 0.86 to 0.74. The second row reports the aggregate growth rate, which increases slightly from 2.8% to 2.89%. Although the number of innovations increases as seen in panel A, most new patents are horizontal innovations. The increase in ϕ decreases the adoption rates of these new patents since they are now mainly sold to monopolists. This generates only a small increase in the aggregate productivity growth rate.

The last row in panel B reports the aggregate welfare in the consumption equivalent, computed by equation (25). We normalize the welfare in the economy before 2008 to be 1. Overall, the tax reduction decreases aggregate welfare by 3%.

Note that in column 2, not only the trade tax decline, but also the firm's income tax. To isolate the effect of the trade tax cut, column 3 fixes ς at 33%, but only decreases the trade tax to 24%. Compared with the first and second columns, we can see I declines, because without a firm income tax reduction, the incentive to invent in I is smaller. Meanwhile, the firm's average Δ is higher, so the composition effect pushes down I.

The horizontal innovation τ increases, as does ϕ, due to the trade tax reduction. This increases the probability of trade, especially inefficient trade. Since the trade becomes very easy, the average

markup increases to 1.14. From panel B, we can see that the misallocation becomes smaller: M only declines to 0.78, because the decrease in I weakens the firm's monopoly power. However, the growth rate is now only 2.78%, which is even lower than the economy before 2008, due to the decline in I.

In column 4, we examine the impact of keeping ϕ the same as the pre-2008 economy while implementing the same tax reform as in column 2. This exercise helps us isolate the impact of ϕ on determining aggregate welfare. We find that both types of innovations, I and τ, increase due to the tax cut. However, because ϕ is fixed, the trade probability and the share of monopoly-power-driven trade remain the same. The adoption rate slightly increases as the reduction in firm income tax makes adoption more attractive. The increase in I and σ^V tends to raise the markup, while the increase in τ and the inventor's adoption rate σ tends to lower the markup. Overall, we find that the average markup is nearly unchanged.

From panel B, we can see that the misallocation M remains the same as in the pre-2008 economy. However, the TFP growth rate is higher, thus leading to a small increase in aggregate welfare. In other words, the welfare cost of the tax reform is mainly due to the change in ϕ.

In summary, while the cut in trade tax can encourage more innovations, it may decrease the aggregate welfare, mainly because the direction of new innovation changes (ϕ is increased) and patents are more likely to be sold to monopolists.

Robustness

The fundamental friction in the model comes from the "monopoly-power-driven" acquisition. To better understand the impact of this friction on the economy, we conduct two sensitivity checks by varying the parameters associated with this fundamental friction.

First, we consider the share of "monopoly-power-driven" acquisition, which is determined by $\beta^V(\phi)$. We vary β_0 to examine its effect on the economy, as shown by the solid line in the left graph of Figure 8. Increasing β_0 will increase the chance of trade and lead to more innovations, but this increase in trade is mainly "monopoly-power-driven" and results in welfare loss. Our calculations suggest that the welfare loss outweighs the gain, leading to a decrease in aggregate welfare when β_0 increases.

We then reduce the firm income tax, as the 2008 tax reform, and plot the welfare for the post-08 economy for each β_0 by the dashed line. The dashed line always lies below the solid line, suggesting that tax reform leads to a welfare loss. Meanwhile, the loss of the tax reform increases with β_0 since the increase in ϕ has a greater impact on the share of "monopoly-power-driven" trade.

Another source of inefficiency in the trade market is that $\sigma^V < 1$, which means that the current monopoly may not be able to utilize the acquired patent efficiently. To investigate the impact of this parameter, we vary α_0 while keeping other parameters unchanged. The solid line in the right graph of Figure 8 plots the aggregate welfare of the pre-08 economy. We observe that increasing α_0 monotonically increases the welfare of the economy. When the monopolist is more likely to adopt the newly acquired technology, the share of "monopoly-power-driven" acquisition decreases. As shown in equation (5), $\alpha(\phi)$ does not enter this equation, so an increase in α_0 results in a smaller distortion relative to the productivity gain, leading to a net increase in aggregate welfare.

The dashed line plots the welfare for the post-08 economy. Over a large range in α_0, the tax reduction generates a welfare loss. However, the loss shrinks as α_0 increases, because when monopolists can utilize patents more efficiently, productivity is more likely to be increased even in "monopoly-power-driven" acquisitions.

Optimal trade tax

We consider three possible thought experiments. To isolate the trade tax, in the following three counterfactuals, we keep the firm income tax rate and individual income tax rate respectively, at 33% and 25%, the levels before 2008. We only change the tax rate when *firms* sell patents. Thus, the numbers are directly comparable to column 3 in Table 7.

The first experiment is focused on determining the optimal flat trade tax rate. The trade-off of the trade tax is as follows. On one hand, a low trade tax can encourage more innovations, particularly those that are highly connected, which can increase the growth rate of the economy. On the other hand, high-connected innovations may have a negative impact on the growth rate. Additionally, a low trade tax can lead to more trade, including inefficient trade. While this can increase the synergy from trade and correct initial misallocations on the invention market, it can also increase the tail of the markup and exacerbate the misallocation effect.

In column 5 of Table 7, we search for the optimal flat firm trade tax rate that maximizes welfare. Our results show that welfare is maximized when $\varsigma = 0.31$, which is approximately 8% higher than the benchmark case (column 1). This finding suggests the model predicts the best tax reduction in 2008 would have been only 2%, which is 6% smaller than the actual tax cut.

The gain from the pro-trade policy in our model depends on the inventor's adoption cost $\bar{d}$, which represents the initial misallocation on the invention market. When this misallocation is high, even the "monopoly-power-driven" trade can improve welfare. To investigate the relationship between the optimal trade tax rate and $\bar{d}$, we conduct a second experiment where we vary

$\bar{d}$ while holding other parameters constant, and find the implied optimal flat trade tax rate as in column 5 of Table 7. Figure 9 shows that the optimal tax rate is a declining function of $\bar{d}$. Therefore, the optimal tax rate depends on the initial misallocation of the invention market. For the Chinese case, when $\bar{d}$ is more than twice the benchmark value, the optimal tax rate is less than 25%.

The fundamental source of inefficiency in our model arises from the two types of acquisitions, which have different implications on productivity and markup. Ideally, if the government could distinguish which trade is "inefficient," and let the trade tax rate be contingent on the types of trade, welfare could be improved. In the last column of Table 7, we consider this possibility, and find that when $\varsigma_{\text{"productivity-driven" trade}} = 0.19$ and $\varsigma_{\text{"monopoly-power-driven" trade}} = 0.37$, the welfare is 43% higher than in the pre-2008 economy. The gain comes from both the increase in M and the increase in the growth rate. In other words, this type of tax raises productivity without distorting the markup.

4.5 Model extension

Up to this point, we have been assuming the quality of new innovations is uniformly distributed, with each innovation improving productivity by one rung on the quality ladder. However, we now introduce a new assumption: in addition to choosing the level of ϕ, the innovator can also select the quality of horizontal innovation, denoted as $q \in \{1, q_H\}$. Therefore, each new patent will now be labeled by both the connection measure ϕ and the quality level q.

If only one buyer is present, the seller will get $E_d(V(q) - w_H d)^+$ (net the trade tax), disregarding who is the buyer.

If two buyers are present, we can have the expected selling price is

$$p(q) = E_{d,d^V,\Delta}[max_{(2)}(V(\Delta + q) - w_H d^V, V(\Delta), V(q), \frac{[V(q) - w_H d]^+}{1 - \varsigma})],$$

where $max_{(2)}$ denotes the second largest element across the members in the bracket. Note that when $q = 1$, $V(\Delta) \geq V(q) > \frac{[V(q) - w_H d]^+}{1-\varsigma}$ since $\Delta \geq 1$. In other words, the monopolist will always buy the low-quality patent. However, when $q = q_H$, the monopolist may not be able to buy the patent, because the best user may offer a price higher than the monopolist.

Then we can see equation (5) changes to

$$\Omega(q, \phi; \varsigma) = (1 - \beta^V(\phi)\beta^X)E_d(V(q) - w_H d)^+ + (1 - \varsigma)\beta^V(\phi)\beta^X p(q). \tag{31}$$

Assume the cost of making a high-quality innovation is $w_H(e(q) + \epsilon(q, \phi))$, where $e(q)$ is

increasing and convex, which is the deterministic component when innovating a patent with quality q. $\epsilon(q,\phi)$ is a random cost component of making an innovation (ϕ,q), following a Type-I Gumbel distribution. Then the share of new patents (ϕ,q) follows

$$\omega(\phi,q)=\frac{exp([\Omega(q,\phi;\varsigma)-e(q,\phi)]/w_H)}{\sum_{\phi',q'}exp([\Omega(q',\phi';\varsigma)-e(q',\phi')]/w_H)}. \tag{32}$$

Note two points. First, when q is large, it is less likely to sell to the monopolist. Hence $p(q)-E_d(V(q)-w_Hd)^+$ is smaller, which means attracting monopolists becomes less attractive for high-quality patents. Hence, high-quality patents may show a smaller ϕ. In the meantime, the tax cut will have a smaller impact for ϕ on high-quality patents.[20]

Second, the gain from selling the patent relative to running the patent by itself $(p(q)-E_d(V(q)-w_Hd)^+)$ is smaller for high-quality patents. When the trade tax declines, the low-quality patents will enjoy more benefits. In other words, a reduction in trade taxes will incentivize the production of more low-quality patents. As a result, the tax reduction could introduce another margin of welfare loss.

To calibrate our new model, we set $e(1)=0$ and choose q_H and $e(q_H)$ to match the production of the top 10% cited patents and the trade probability of the top 10% cited patents relative to the average traded probability of a patent. We keep all other parameters the same as reported in Table 5. We then consider three different scenarios: an economy with the same tax rate as pre-2008, an economy with the same tax rate as post-2008, and an economy where only the trade tax rate has been reduced. The results of our analysis are presented in Table 9.

Our results indicate that the increase in ϕ and the decline in q resulting from the tax reduction can lead to a welfare loss, due to the increase in "monopoly-power-driven" trade, which is a robust finding in our model. Specifically, we find that welfare declined by 2% after 2008, and if we only reduce the trade tax, the welfare loss is even greater.

5 Conclusion

The paper utilizes the China tax reform to document the causal impacts of pro-patent trade subsidies. The cut on the patent trade tax increases the number of new innovations but increases the connection of new innovations with big firms. These patterns result in a large increase in "monopoly-power driven" patent trade.

To understand the macro consequence of the pro-patent trade policy, we build an endoge-

[20]This prediction is consistent with the empirical observation in Figure 3, where the citation-weighted average ϕ shows a stronger increase than the equal-weighted average ϕ.

nous growth model in which the technology advantages determine firms' monopoly powers on the product market. Patent trade can change the incumbents' technology advantages hence their monopoly powers. Meanwhile, new innovators will also change their innovation decisions to maximize the gain from trade. Specifically, they will innovate more patents to maintain incumbents' high markups, but fewer patents to improve real productivity.

In the context of China's tax reform, we find the cut in the patent trade tax decreases the overall welfare mainly because of the distortion from higher market concentration.

References

Acemoglu, D. and U. Akcigit (2012). Intellectual property rights policy, competition and innovation. *Journal of the European Economic Association 10*(1), 1–42.

Acemoglu, D., U. Akcigit, H. Alp, N. Bloom, and W. Kerr (2018, November). Innovation, reallocation, and growth. *American Economic Review 108*(11), 3450–91.

Akcigit, U., M. A. Celik, and J. Greenwood (2016). Buy, keep, or sell: Economic growth and the market for ideas. *Econometrica 84*(3), 943–984.

Arkolakis, C., A. Costinot, D. Donaldson, and A. Rodríguez-Clare (2019). The elusive procompetitive effects of trade. *The Review of Economic Studies 86*(1), 46–80.

Bessen, J. and M. J. Meurer (2008). Do patents perform like property? *Academy of Management Perspectives 22*(3), 8–20.

Bloom, N., R. Griffith, and J. Van Reenen (2002). Do r&d tax credits work? evidence from a panel of countries 1979–1997. *Journal of Public Economics 85*(1), 1–31.

Bloom, N., J. Van Reenen, and H. Williams (2019). A toolkit of policies to promote innovation. *Journal of Economic Perspectives 33*(3), 163–84.

Bryan, K. A. and E. Hovenkamp (2020). Antitrust limits on startup acquisitions. *Review of Industrial Organization 56*, 615–636.

De Ridder, M. (2019). Market power and innovation in the intangible economy.

Edmond, C., V. Midrigan, and D. Y. Xu (2015). Competition, markups, and the gains from international trade. *American Economic Review 105*(10), 3183–3221.

Edmond, C., V. Midrigan, and D. Y. Xu (2018). How costly are markups? Technical report, National Bureau of Economic Research.

Gans, J. S. and S. Stern (2000). Incumbency and r&d incentives: Licensing the gale of creative destruction. *Journal of Economics & Management Strategy 9*(4), 485–511.

Hall, B. H. and M. MacGarvie (2010). The private value of software patents. *Research Policy 39*(7), 994–1009.

Hopenhayn, H., J. Neira, and R. Singhania (2018). The rise and fall of labor force growth: Implications for firm demographics and aggregate trends. *NBER Working Paper*, 1–28.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics 124*(4), 1403–1448.

Jones, C. I. (2022). The end of economic growth? unintended consequences of a declining population. *American Economic Review 112*(11), 3489–3527.

Klette, T. J. and S. Kortum (2004). Innovating firms and aggregate innovation. *Journal of political economy 112*(5), 986–1018.

König, M., K. Storesletten, Z. Song, and F. Zilibotti (2022). From imitation to innovation: Where is all that chinese r&d going? *Econometrica 90*(4), 1615–1654.

Kroen, T., E. Liu, A. R. Mian, and A. Sufi (2021). Falling rates and rising superstars. Technical report, National Bureau of Economic Research.

Lashkaripour, A. and V. Lugovskyy (2018). Scale economies and the structure of trade and industrial policy. Technical report, mimeo Indiana University.

Lucas Jr, R. E. and B. Moll (2014). Knowledge growth and the allocation of time. *Journal of Political Economy 122*(1), 1–51.

Olley, G. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Olmstead-Rumsey, J. (2019). Market concentration and the productivity slowdown.

Perla, J. and C. Tonetti (2014). Equilibrium imitation and growth. *Journal of Political Economy 122*(1), 52–76.

Peters, M. (2020). Heterogeneous markups, growth, and endogenous misallocation. *Econometrica 88*(5), 2037–2073.

Serrano, C. J. (2010). The dynamics of the transfer and renewal of patents. *The RAND Journal of Economics 41*(4), 686–708.

Serrano, C. J. (2018). Estimating the gains from trade in the market for patent rights. *International Economic Review 59*(4), 1877–1904.

Sidak, J. G. (2007). Holdup, royalty stacking, and the presumption of injunctive relief for patent infringement: a reply to lemley and shapiro. *Minn. L. Rev. 92*, 714.

Wei, S.-j., J. Xu, G. Yi0, and X. Zhang (2022). Mild government failure. *Working paper*.

Tables and Figures

Table 1: Shares of Corporate Patents across Industries

Industries	Share of corporate patents	Economy of scale (Lashkaripour and Lugovskyy (2018))
Petroleum	0.69	1.98
Tobacco	0.61	-
Transportation equipment	0.59	0.58
Electrical machinery and equipment	0.54	1.37
Chemical products	0.53	0.36
Special equipment	0.51	0.32
General equipment	0.51	0.32
Metal	0.50	0.26
Pharmaceutical manufacture	0.46	0.36
Instrumentation and office machinery	0.43	0.32
Food and beverage	0.41	0.42
Communication equipment and computer	0.40	0.32
Textile, apparel, shoes and leather	0.37	0.28
Non-metallic mineral products	0.36	0.22
Fiber, rubber and plastic	0.32	0.24
Wood and furniture	0.29	0.34
Paper products	0.24	0.18
Handicrafts	0.21	-
Printing, publishing, educational and sporting goods	0.20	-

Notes: This table shows shares of corporate patents and estimates of economy of scale by Lashkaripour and Lugovskyy (2018) across industries.

Table 2: Shares of Corporate Patents across Industries: Industry Component

	Ind. FE	Yr. FE
R-squared	0.67	0.165
Adj R-squared	0.60	0.167

Notes: This table shows the R-square value when regressing the industrial share of corporate patents on industry dummies and year dummies, respectively.

Table 3: Effects of Patent Trade on Industry Productivity, Size, and Markup

Dependent var	(1) pre & post	(2) pre & post	(3) pre & post	(4) pre & post	(5) pre only
ln(TFP)					
	0.040	0.014	0.051*	0.075***	-0.037
	[0.032]	[0.039]	[0.033]	[0.027]	[0.091]
Ave. markup					
	0.042*	0.025**	0.105*	0.107*	0.003
	[0.024]	[0.021]	[0.052]	[0.058]	[0.002]
Herfindahl index					
	0.040**	0.092***	0.072***	0.063*	-0.011
	[0.021]	[0.016]	[0.023]	[0.036]	[0.042]
IV	N	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y
Exclude 8 ind			Y		
Observed industry char				Y	
Obs.	872	872	592	872	124

Notes: Each cell in this table is a regression. All specifications are as equation (3) with different dependent variables. Observed industry characteristics include the sales share of SOEs and average sales per firm in each year. Standard errors (in brackets) are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Effects of Patent Trade on New Innovations

Dependent var	(1) pre & post	(2) pre & post	(3) pre & post	(4) pre & post	(5) pre & post	(6) pre only
ln(New patent count)						
	0.479**	0.683*	1.002**	0.583***	0.666**	0.017
	[0.236]	[0.373]	[0.409]	[0.216]	[0.301]	[0.302]
ln(Connection)						
	0.111	0.220**	0.378**	0.181**	0.253***	-0.031
	[0.060]	[0.051]	[0.189]	[0.080]	[0.072]	[0.100]
IV		Y	Y	Y	Y	Y
Industry FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Corporate inventor only			Y			
Exclude 8 ind				Y		
Observed industry char					Y	
Obs.	872	872	872	592	872	124

Notes: Each cell in this table is a regression. All specifications are as equation (3) with different dependent variables. Observed industry characteristics include the sales share of SOEs and average sales per firm in each year. Standard errors (in brackets) are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Parameter Values Used in the Baseline Simulation

Parameter	Value	Description	Source
c^V	0.56	In-house vertical innovation cost	Calibration
c^X	0.58	In-house horizontal innovation cost	Calibration
v^E	0.18	Entry cost of individual inventor	Calibration
β_0	0.17	High connection patents' trade prob.	Calibration
β_1	0.17	Low connection patents' trade prob.	Calibration
β^X	0.01	Productivity-driven trade prob.	Calibration
$\bar{d}$	0.43	Inventor's adoption cost	Calibration
α_0	0.98	High connection patents' adoption cost	Calibration
α_1	0.25	Low connection patents' adoption cost	Calibration
λ	1.06	Step size	Calibration
H	0.33	Number of researchers	Calibration
L	1	Number of workers	Normalization
ρ	0.05	Discount rate	Imposed

Table 6: Model Fitness

Targeted Moments	Data	Model
TFP growth(%)	2.80	2.80
$\frac{w_H}{w}$	1.54	1.54
Entry rate of firms	0.05	0.05
Entry rate of patents	0.11	0.11
Markup of 5 years old firm/new entrant	1.15	1.15
Sales of 5 years old firm/new entrant	1.08	1.08
Trade prob: 25th percentile ϕ	0.02	0.02
Markup change aft. acquisition: 25th percentile ϕ	0.01	0.01
Sales change aft. acquisition: 25th percentile ϕ	0.01	0.01
Trade prob: 75th percentile ϕ	0.04	0.04
Markup change aft. acquisition: 75th percentile ϕ	0.07	0.07
Sales change aft. acquisition: 75th percentile ϕ	0.03	0.03

Notes: This table reports the model's fit by comparing the moments in the model and the data.

Table 7: Welfare Effects of Tax Cut

	(1) Bef 2008	(2) Aft 2008	(3) Only trade tax cut	(4) Shut down change in ϕ	(5) Best flat tax	(6) Optimal tax
			A: Endogenous outcomes			
I	0.56	0.59	0.54	0.57	0.56	0.57
τ	0.24	0.35	0.31	0.25	0.28	0.27
Ave. ϕ	0.29	0.39	0.33	0.29	0.31	0.25
Trade prob of a new patent	0.14	0.28	0.21	0.14	0.18	0.12
Monopoly-power-driven trade share	0.27	0.45	0.42	0.27	0.31	0.17
σ^V	0.13	0.19	0.14	0.14	0.15	0.12
Ave adoption rate	0.48	0.42	0.41	0.49	0.47	0.53
Ave. markup	1.11	1.17	1.14	1.11	1.09	1.06
			B: Welfare decomposition			
M	0.86	0.74	0.78	0.86	0.89	0.96
TFP growth (%)	2.80	2.89	2.78	2.81	2.92	3.37
Welfare (C%)	1.00	0.97	0.94	1.01	1.08	1.43

Notes: This table reports the impacts of tax cuts in different scenarios. Column 1 is the benchmark economy before 2008. In column 2, the firm's income tax declines from 33% to 25%, while the trade tax declines from 30% to 24%. In column 3, we only lower the trade tax, and the income tax remains at 33%. In columns 4 and 5, we seek optimal trade tax rates while decreasing the firm's income tax rate to 25%. In column 4, we seek the best flat trade tax rate to maximize welfare. In column 5, the trade tax rate can depend on trade types (i.e, "productivity-driven" acquisition vs "monopoly-power-driven" acquisition).

Table 8: Change in Firm Markup after Acquisition: Data vs Simulated Sample

Dependent var. $d\ln\mu_{f,t}$	(1) Actual data	(2) Simulated sample
$\ln\mu_{f,t0} \times Post_t$	0.117**	0.141***
	[0.042]	[0.035]
$\ln\mu_{f,t0}$	-0.016	-0.225***
	[0.053]	[[0.048]
$Post_t$	0.054*	0.119***
	[0.022]	[0.037]
$\ln sale_{f,t-1}$	-0.136*	-0.222***
	[0.069]	[0.017]
Other controls	Y	
Adj R2	0.23	0.54
Obs.	5,743	5,743

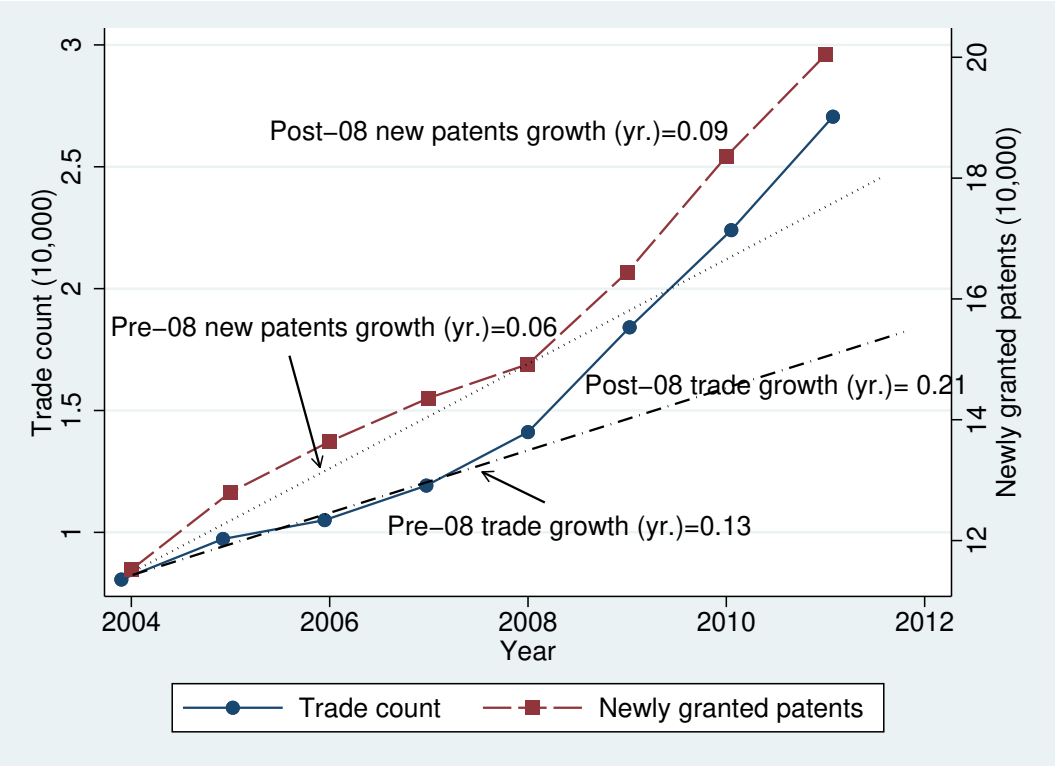
Notes: This table compares the change in firm markup after acquisition (measured by two years before and after the acquisition). The first column shows the result using the actual data and the second column uses the model-simulated data. $Post_t = 1$ if year $t \geq 2008$ and $\ln\mu_{f,t0}$ in the firm markup in 2004. Other controls include the buyer's industry, located city, and three-years-forward citations of the acquired patents. Standard errors (in brackets) are clustered at the industry-year level. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Welfare Effects of Tax Cut: Extension Model

	(1) Bef 2008	(2) Aft 2008	(3) Only trade tax cut
	A: Endogenous outcomes		
I	0.57	0.68	0.56
τ	0.32	0.43	0.35
Ave. ϕ	0.36	0.41	0.38
Ave quality of new patents	1.19	1.20	1.07
Trade prob of a new patent	0.17	0.30	0.27
Monopoly-power-driven trade share	0.32	0.45	0.45
σ^V	0.15	0.22	0.23
Ave adoption rate	0.49	0.43	0.44
Ave. markup	1.14	1.22	1.19
	B: Welfare decomposition		
M	0.88	0.79	0.80
TFP growth (%)	2.86	2.92	2.87
Welfare (C%)	1.00	0.98	0.96

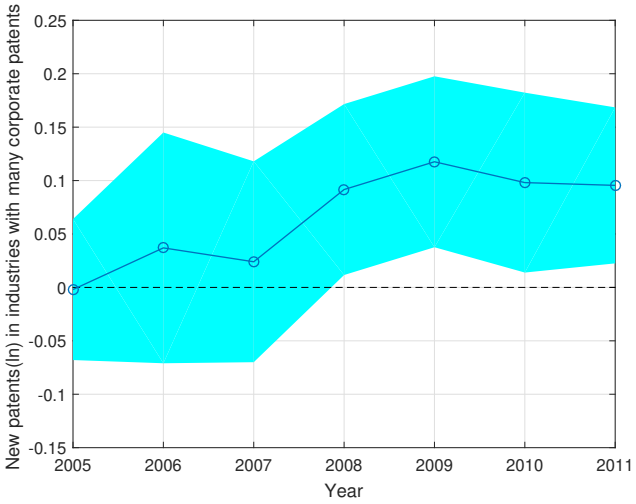
Notes: This table reports the impacts of tax cuts in different scenarios of the extension model, in which the quality of the horizontal innovation is optimally chosen. Column 1 is the economy before 2008. In column 2, the firm's income tax declines from 33% to 25%, while the trade tax declines from 30% to 24%. In column 3, we only decrease the trade tax, and the income tax remains at 33%.

Figure 1: Patent trade and newly granted patents from 2004 to 2011

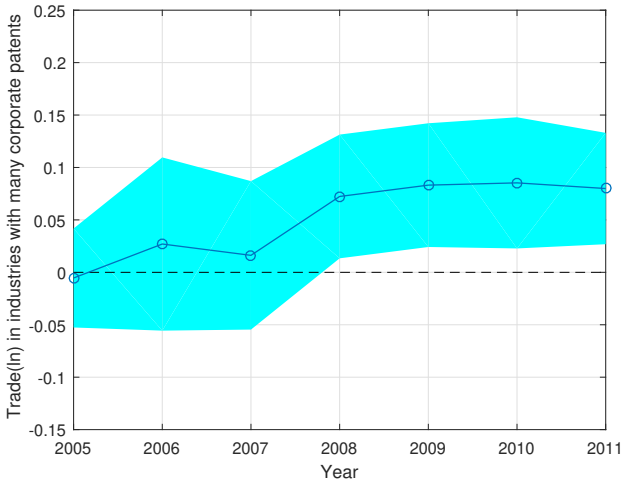


Notes: This figure shows the number of trade (solid-dot line, left y-axis) and newly granted patents (square-dashed line, right y-axis).

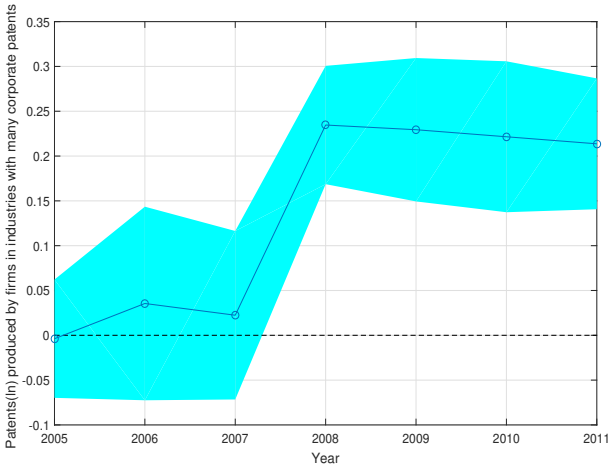
Figure 2: Patent production and trade rise especially in industries with many corporate patents



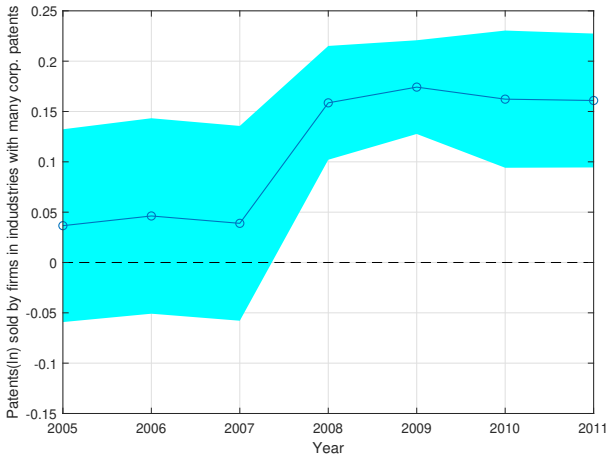
ln(Patents)



ln(Trade)



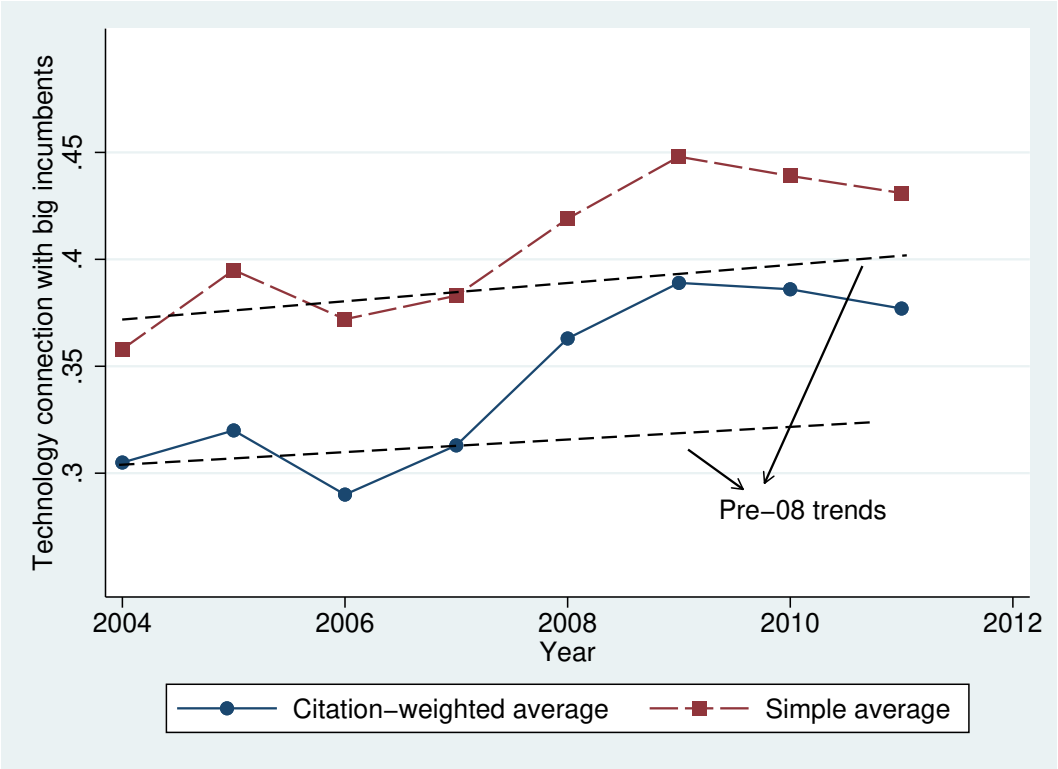
ln(Patents produced by firms)



ln(Patents sold by firms)

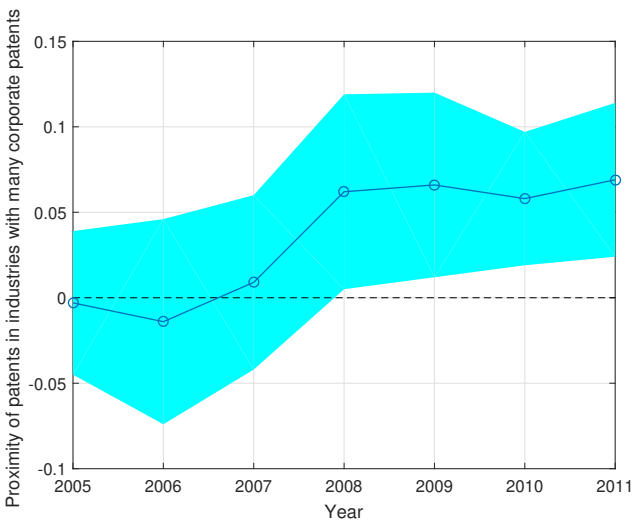
Notes: This figure plots the change in the citation-weighted count of patents produced or sold when the industry's share of corporate patents in 2004 is 1% higher (β_t from equation (3)). The grey areas represent 95% confidence intervals.

Figure 3: Average of the proximity of new patents with big incumbents' technology from 2004-2011

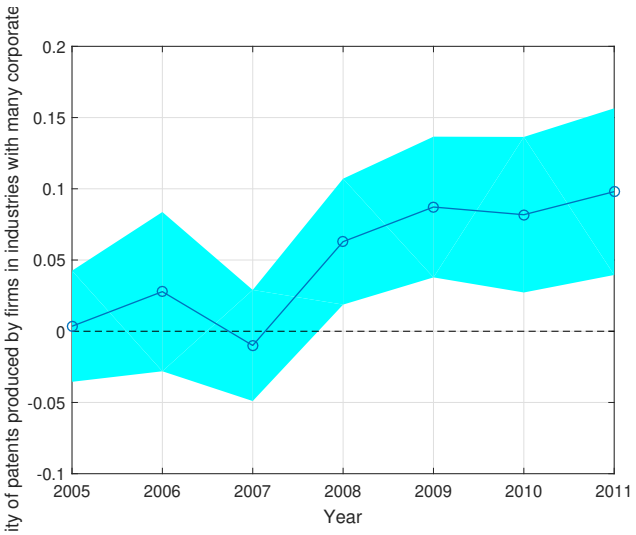


Notes: This figure shows the average of ϕ of newly granted patents.

Figure 4: Proximity of new patents with big incumbents' technology rises especially in industries with many corporate patents



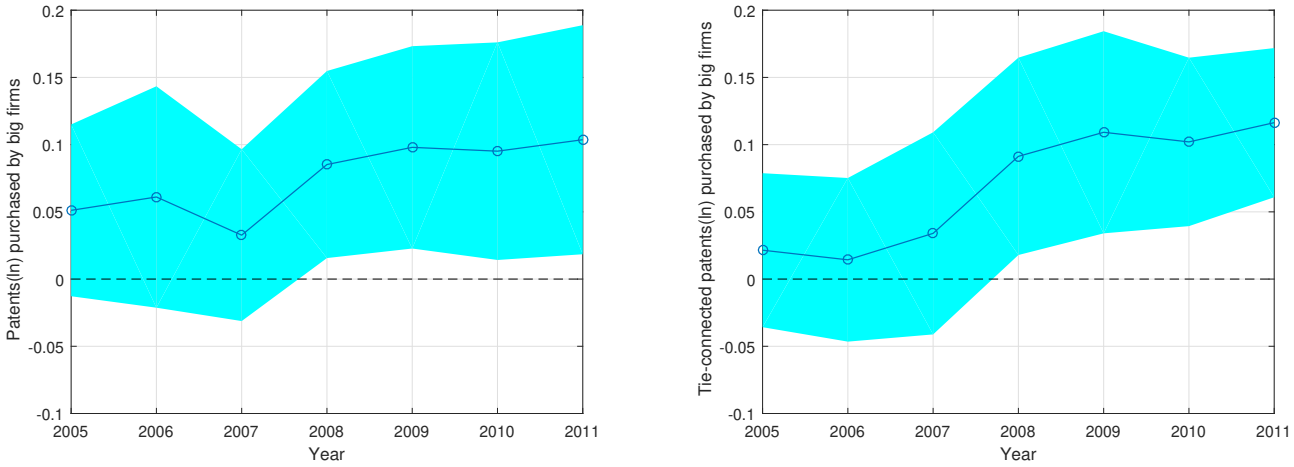
ϕ: all patents



ϕ: patents produced by firms

Notes: This figure plots the change in proximity of new patents with big incumbents' technology when the industry's share of corporate patents in 2004 is 1% higher. The grey areas represent 95% confidence intervals.

Figure 5: Tie-connected patents purchased by big buyers rise in industries with many corporate patents



ln(Patents purchased by big firms) ln(Tie-connected patents purchased by big firms)

Notes: This figure plots the change in trade count and patents sold by firms when the industry's share of corporate patents in 2004 is 1% higher (β_t from equation (3)). The grey areas represent 95% confidence intervals.

Figure 6: Expansions of the firm

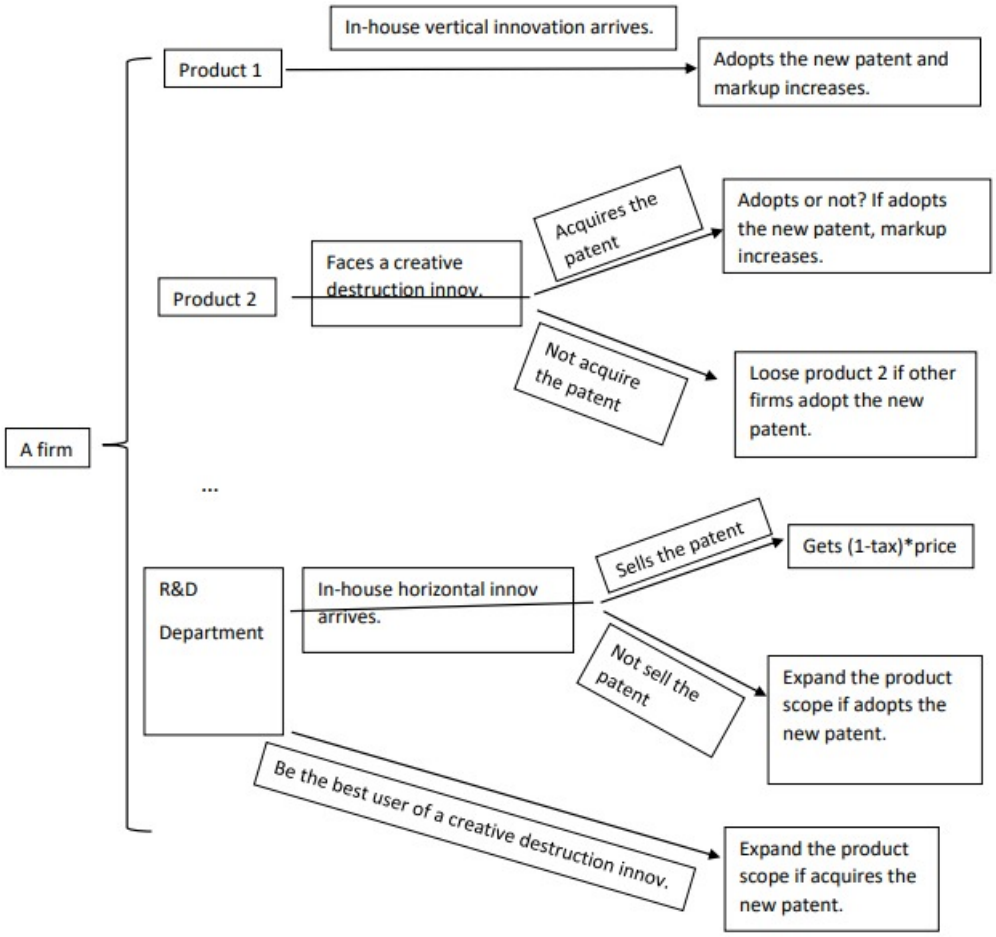
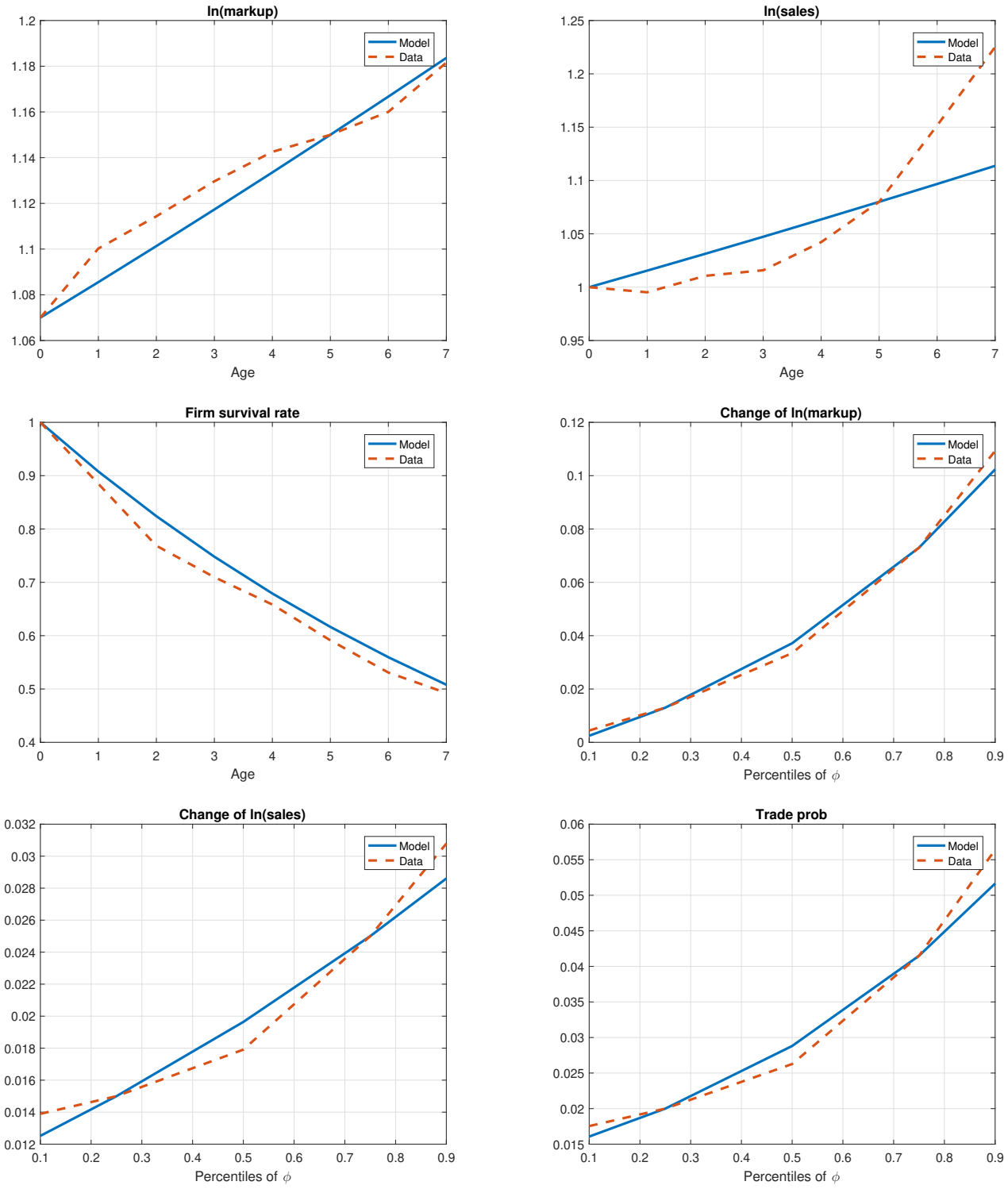


Figure 7: Non-targeted moments: Model vs Data



Notes: The first three graphs compare the model and the data for the life-cycle of markups, sales growth, and survival rate. The next three graphs compare the model and the data for the change in markups and sales after the acquisition, and the traded probability conditional on the patent's connection with current technology, ϕ.

Figure 8: The impact of varying the probability of monopoly-power-driven trade β_0 and adoption cost of monopolists α_0 on welfare of pre- and post-08 economies

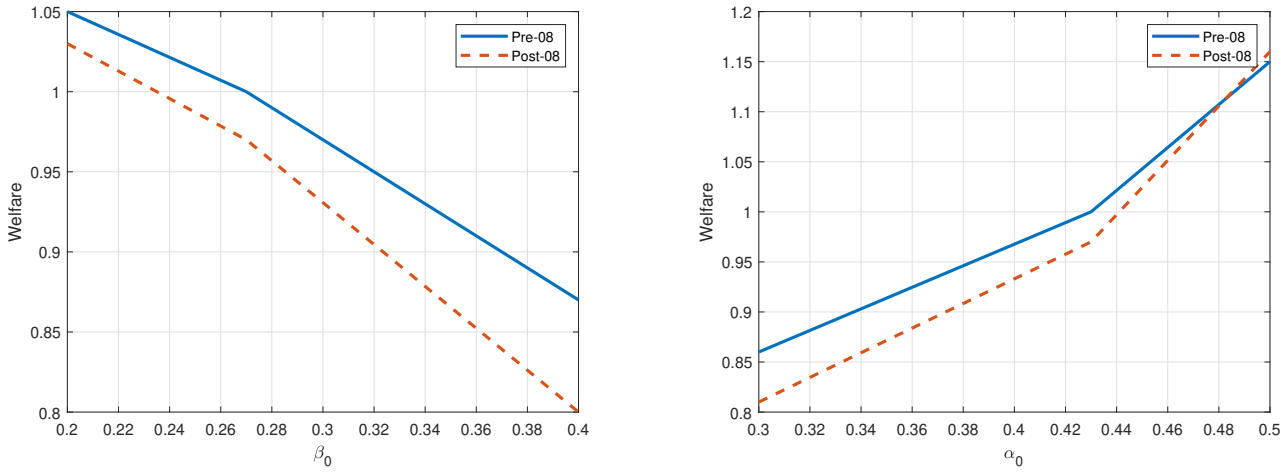
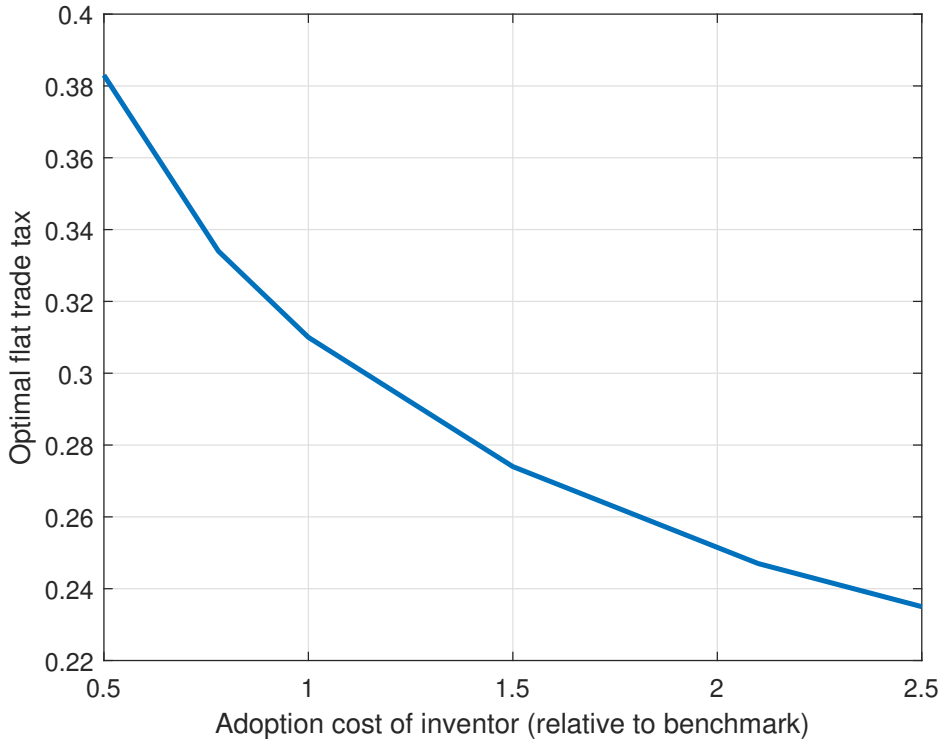


Figure 9: Optimal trade tax under different adoption cost of inventor $\bar{d}$



Online Appendix (not for publication in print)

A Proof of proposition 1

From equation (11), we can see x does not depend on Δ, since Γ is a constant on the BGP. Meanwhile, from equations (6) and (9), θ^V and θ^X are two constants.

When $b = 1$, we can guess

$$\hat{V}(\Delta) = A - B\lambda^{-\Delta},$$

where $\hat{V}(\Delta) = \frac{V(\Delta)}{w}$.

Substituting this guess into (11), we can have

$$I = \left[\frac{B\left(1-\lambda^{-1}\right)}{\frac{w_H}{w}v^I}\right]^{\frac{1}{\zeta}}$$

$$x = \left[\frac{\Gamma}{\frac{w_H}{w}v^X}\right]^{\frac{1}{\zeta}}.$$

From (6), we can have

$$\hat{B}^V(\Delta) = \lambda^{-\Delta}Q_1 + A - Q_2,$$

where $\hat{B}^V = \frac{B^V}{w}$. Q_1 and Q_2 are two constants. $Q_1 = E_{\phi,\tilde{d}^V}\left[\beta^V(\phi)\left[B\left(1-\lambda^{-1}\right) - \tilde{d}^V\right]^+ - B\right]$ and $Q_2 = E_\phi \beta^V(\phi)\left[\beta^X \frac{\left[\hat{V}(1)-d\right]^+}{1-\zeta} + \left(1-\beta^X\right)\hat{V}(1)\right]$.

Hence the Bellman equation implies that A and B satisfy

$$(\rho+g)A = (1-\zeta)\frac{Y}{w} + x\Gamma - \frac{w_H}{w}\frac{v^X}{1+\zeta}x^{1+\zeta} - \tau Q_2 + \tau B^X$$

$$(\rho+g)B = (1-\zeta)\frac{Y}{w} - IB\left(1-\lambda^{-1}\right) + \frac{w_H}{w}\frac{v^I}{1+\zeta}I^{1+\zeta} - \tau\left(Q_0 - B\right),$$

and

$$\sigma^V = \Pr\left[\tilde{d}^V < B\left(1-\lambda^{-1}\right)\right].$$

B Proof of proposition 3

First, consider the special case of $b = 1$. In a stationary equilibrium, from equation (28), we have

$$h(\Delta) = \left(\frac{1}{1+k}\right)^{\Delta} k,$$

where $k = \frac{\hat{\tau}}{I^* + \tau\theta^V \sigma^{V*}}$. Hence $\Pr\left(\Delta > \bar{\Delta}\right) = e^{-\bar{\Delta}\ln(1+k)}$. Thus, the markup distribution satisfies

$$\begin{aligned} \Pr\left(\lambda^{\Delta} > \bar{\mu}\right) &= \Pr\left(\Delta > \ln\bar{\mu}/\ln\lambda\right) \\ &= \bar{\mu}^{-\chi}. \end{aligned}$$

If b does not equal 1, we know that $I(\Delta)$ and $\sigma^V(\Delta)$ will equal to I^* and σ^{V*} when Δ is big enough. Hence, the markup still has a Pareto tail.