What Can Explain the Chinese Patent Explosion?

By Markus Eberhardt,[a] Christian Helmers,[b] and Zhihong Yu[c]

[a] School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD and

Centre for the Study of African Economies, University of Oxford; e-mail: Markus.Eberhardt@nottingham.ac.uk

[b] Santa Clara University, and Centre for the Study of African Economies, University of Oxford

[c] School of Economics, University of Nottingham

Abstract: We analyse the 'explosion' of patent filings by Chinese residents both domestically and in the United States during the early 2000s, employing a unique dataset of 374,000 firms matching patent applications to manufacturing census data. Our analysis reveals that patenting is highly concentrated among a small number of firms, operating in the information and communication technology sector. Although increases in patent filings by these companies are partly driven by increased R&D intensity, our analysis suggests that the explosion of patent filings at the Chinese patent office is driven by factors other than underlying innovative behavior, including government subsidies that encourage patent filings directly.

JEL classification: L25, O12

1   Introduction

China's economic success over the past decades has been widely regarded as the result of its ability to produce manufactured goods at low cost, building on the availability of cheap labour and scale economies, while relying on existing technologies of production. China's ability to upgrade its technology base and move up the value-chain is frequently argued to be hampered by weak (intellectual) property rights enforcement (Zhao, 2006). More recently, however, the notion that China is catching up fast in terms of scientific and technological innovation has gained considerable ground. The number of domestic invention patent filings with the Chinese patent office (SIPO) has increased at an average rate of 32% per annum from around 15,600 to over 700,000 during the period 1999-2013.[1] Utility patent filings by Chinese residents[2] with the U.S. patent office (USPTO) grew at an annual rate of 35% to nearly 15,500 over the same period, albeit from a low base of 271 in 1999.[3] This patent 'explosion' at home and abroad is paired with strengthened statutory intellectual property (IP) rights protection (Park, 2008) and an increased interest by policymakers in the role of IP in fueling domestic innovation, with a particular focus on foreign technology transfer and incentives to invest in R&D. Accordingly, the recent National Patent Development Strategy (2011-2020) envisions an increase in the number of annual patent applications (including invention, utility patents and designs) from 1.2 million in 2010 to 2 million in 2015. The plan also foresees a doubling of the number of patent applications filed by Chinese applicants abroad over the same time horizon. These ambitious targets reflect a positive outlook in parts of the literature on Chinese innovation, the Chinese IP rights system, and Chinese development in general (Fischer and von Zedtwitz, 2004; Subramanian, 2011).

At the same time, there is some evidence to suggest that most of the innovation in China is of merely incremental nature and hence the corresponding patents protect 'small inventive steps' (Puga and Trefler, 2010). While such incremental innovation may still be valuable and in fact account in large part for China's success (Breznitz and Murphree, 2011), the concern is that the recent increase in patent applications is produced overwhelmingly by inventions embodying little technological progress. Recent empirical evidence suggests that patent subsidies, introduced by local governments in virtually all Chinese provinces from 1999 onwards, have also played an important role in explaining the 'explosive' growth of Chinese patenting (Li, 2012; Dang and Motohashi, 2015). Boeing and Mueller (2015) sug-

---

[1]Data from the World Intellectual Property Organization (WIPO).

[2]Throughout the paper we use 'Chinese firms' and 'Chinese residents' interchangeably. Our firm-level dataset covers indigenous firms as well as subsidiaries of foreign multinationals. U.S. utility patents correspond to invention patents in China.

[3]Data taken from various USPTO Performance and Accountability Reports.

gest that patent quality of PCT filings[4] by Chinese applicants is low by international comparison and that quality has been decreasing over time as the number of filings has increased. They also find some evidence for a negative correlation between patent quality and filing subsidies.

The view that China's patent explosion over the past two decades was driven largely by an increase in the patenting of low quality inventions — fueled by public incentive schemes — stands in stark contrast to earlier findings in the literature, which explained the recent increase in Chinese firms' patenting activity by an influx of FDI, the opening of the economy in particular through China's WTO accession, and a major overhaul of the legal framework in form of amendments of the patent law (Hu and Jefferson, 2009). Despite widespread doubts about the link between innovative prowess and the Chinese patent explosion in the media and in policy circles,[5] there is no quantitative analysis based on representative firm-level data that investigates the determinants of the Chinese patent explosion during its critical years in the early 2000s.

We analyze the recent 'explosion' in the number of patent applications by manufacturing firms registered in China with SIPO as well as the USPTO, which is by far the most important destination for Chinese patent filings abroad (Wunsch-Vincent et al., 2015). In contrast to the study by Hu and Jefferson (2009) our analysis is focused on 'invention' patents which are subject to substantive examination for novelty and inventiveness in both constituencies; this prevents our analysis from being distorted by the vast number of utility models and design patents with low innovative content that do not require substantive examination by the Chinese or U.S. patent offices. Apart from separately analysing the determinants of patenting with SIPO and the USPTO, we infer information on underlying inventions by assessing *where* companies seek patent protection: only domestically with SIPO or (also) with the USPTO. Not only are the direct and indirect costs associated higher in the U.S., but inventions are required to overcome a higher novelty hurdle in patent examination during our sample period. These differences suggest that a comparison of patents filed with the USPTO and SIPO reveals additional information on the underlying invention and the corresponding patentees.

We construct a representative firm-level dataset that combines invention patent data and company financials. We match SIPO and USPTO patents filed between 1985 and 2006 to around 316,000 manufacturing firms contained in China's Annual Survey of Industrial Enterprises (ASIE) compiled by the National

---

[4]Filings under the 'Patent Cooperation Treaty' allow an inventor to simultaneously seek protection in a large number of countries using a single application.

[5]In particular *The Economist* magazine has voiced repeated concerns that 'merely churning out patents does little to advance innovation' (Dec 13th 2014; see also Oct 14th 2010).

Bureau of Statistics of China (NBS) for the period 1999-2006.[6] The period covered represents perhaps the most interesting period in state innovation and IP policy as well as firm innovation activity in China: it encompasses aggressive opening up to FDI, policy commitments related to WTO-entry in 2001, a substantial increase in exporting, an amendment to the patent law, increased government incentives to patent and an accelerated pace of privatisation (Fischer and von Zedtwitz, 2004; Naughton, 2007; Hu and Mathews, 2008; Li, 2012; Dang and Motohashi, 2015).

Our descriptive analysis shows that a small number of Chinese companies concentrated in the information and communication technology (ICT) equipment industry accounts for a large share of the dramatic increase in USPTO patents held by Chinese residents, with underlying technologies mostly related to electronics and semiconductors. This select group of firms also accounts for a substantial share of SIPO patents though there is a larger number of companies across a wider range of industries obtaining domestic patent protection. The concentration of patenting in an industry that has moved from 'Patent Portfolio Races' during the 1990s (Hall and Ziedonis, 2001) to outright 'Patent Wars' (*Financial Times*, 17th October 2011) casts some doubt on the underlying technological value of the steep increase in patent counts produced by Chinese firms in this sector. Previous empirical work on Chinese patenting missed this concentration since analysis was based on aggregate data (Sun, 2000; Hu and Mathews, 2008; Hu, 2010) or self-reported patenting without distinction between low-sophistication design or utility and more substantive invention patents (Hu and Jefferson, 2009). Comparing the descriptive statistics for patenting with non-patenting firms, and for those firms patenting in the U.S. with those exclusively patenting in China, reveals a large number of significant differences to motivate our empirical analysis.

We rely on the patent production function approach (Pakes and Griliches, 1980; Hall and Ziedonis, 2001) to explain the patenting decision and number of patent filings by Chinese companies with SIPO and the USPTO, respectively. Apart from the standard predictors of patenting, such as R&D expenditure, firm size, and age, we are particularly interested in the importance of a firm's exporting behavior, financial constraints, as well as province-level patent subsidies in predicting patenting behaviour. There is a large literature showing a positive effect of innovation on exporting (Salomon and Shaver, 2005; Lachenmaier and Woessmann, 2006; Girma et al., 2008; Harris and Li, 2009; Ganotakis and Love, 2011; Melitz and Trefler, 2012), which suggests that exporting in turn should predict patenting provided the patents reflect underlying innovations. Similarly, financial variables are key determinants of corporate innovation activities (Brown et al., 2009, 2012; Guariglia and Liu, 2014) and may help identify structural differences

---

[6]Our regressions also include firms which are not part of our Qin/Oriana bridge dataset (see Section 2): we empirically account for selection from the larger ASIE (374,000 firms) into the integrated ASIE-Qin/Oriana (316,000 firms) dataset.

between types of firms based on where they chose to safeguard their IP rights. Finally, with specific reference to China there is recent evidence which suggests that state subsidies are an important element in explaining patent filings of Chinese firms (Li, 2012; Dang and Motohashi, 2015; Lei et al., 2015) and we add information on provincial patent filing subsidy schemes to our patent production functions.

Our findings confirm that patent filings with SIPO are in part driven by state incentive schemes, and we further document a negative correlation between export intensity and domestic patenting. In contrast, for USPTO patentees resident in China the incentive variable is insignificant and export intensity is positively correlated with foreign patenting. Those companies in China filing with the USPTO are substantially larger in terms of number of workers than those only filing domestically. Financial constraints play an important role in innovation behaviour but do not appear to be a source of differential firm behaviour eliciting qualitative differences. Our findings thus suggest that domestic patenting in China, on average, is driven by state incentives and distinct 'types' of firms (in terms of size and export intensity) compared with those firms patenting overseas with the USPTO.

Our analysis contributes to the literature on innovation and economic development (Nordhaus, 1969; Penrose, 1973) by exploring the drivers behind a dramatic shift in the number of patent filings in China. Our results illustrate that large increases in domestic patenting activity *per se* cannot be seen as indicative of associated changes in innovative behavior in a developing country context. The strong concentration of patenting in ICT that we find in China on the one hand, and the impact of public incentive programs as well as the inverse export-patenting relationship on the other, further caution that a broader technological take-off is not (yet) occurring. That said, other successful Asian economies have seen similar concentrations in patenting activity, in particular during the early take-off phases.[7]

The remainder of this paper is organized as follows. Section 2 discusses the construction of our dataset. Section 3 explains our empirical strategy. Sections 4 and 5 discuss some descriptive evidence and our analytical results. Section 6 offers some brief concluding thoughts.

---

[7]Mahmood and Singh (2003) point to a strong concentration of USPTO patents (1970-1999) among assignees in South Korea and Singapore as the top 50 assignees hold 85% and 70% of each country's USPTO patents, respectively.

2   Data

2.1   Firm-level Data

Our firm-level data come from China's Annual Survey of Industrial Enterprises (ASIE) compiled by the National Bureau of Statistics of China (NBS). ASIE includes the whole population of state-owned firms as well as all non-state-owned companies with annual sales above CNY5 million (around US$600,000). On average, over 200,000 firms from all regions of China are included each year, accounting for 95% of total Chinese industrial output and 98% of industrial exports, covering 39 two-digit industries, of which 30 belong to manufacturing industries. Our data cover four distinct years in the period 2001-6, with a sample of over 800,000 observations from 374,000 firms. Key variables include a unique firm identifier, R&D expenditure (representing the binding constraint for analysis: in the version of ASIE available to us this variable is only reported in 2001, 2002, 2005, and 2006), exports, ownership, output, sales, employment, and industry of operation.[8]

2.2   Patent Data

The patent data come from the European Patent Office's PATSTAT database (version 10/2010). We extract patents filed by Chinese residents (this includes indigenous and foreign(-invested) firms). Our analysis focuses on the application date of a patent. We obtain information on the grant status of patent filings from a 2014 version of PATSTAT to account for a grant lag of several years.

2.3   Matching/Bridge

Due to the absence of a unique identifier shared by the firm-level and patent data, the main data problem consists in matching patents to firms. This is generally challenging for a number of reasons (Helmers et al., 2011); in the case of China, matching is even more difficult due to the different ways in which firm names can be recorded: using (a) Chinese characters, (b) pinyin transcription, (c) a translation of the Chinese names into English, and (d) any mix of (a)-(c).

The Chinese census data contain only firm names using Chinese characters (a), whereas PATSTAT contains (b), (c) and (d). In principle, to match patents to firms we would have to either transcribe firms'

---

[8]In line with the existing literature (e.g. Guariglia and Liu, 2014), we exclude observations with negative values of output, sales, exports, capital or intermediate inputs; and further observations with total assets less than total fixed assets or total liquid assets or with total sales less than exports.

names contained in ASIE or the assignee names in PATSTAT. Instead we identified an alternative so-lution: the Qin and Oriana databases provided by Bureau Van Dijk offer firm-level balance sheet data for individual firms in the Asia-Pacific region. The combination of Qin/Oriana contain data for about 451,000 Chinese firms for 2001-2009. The advantage of using Qin/Oriana is that these report firm names using the Latin alphabet as well as the ASIE unique firm identifier. This allows us to link Qin/Oriana to ASIE through the unique identifier and to use Qin/Oriana firm names to match with assignee names contained in PATSTAT. While this approach allows us to match patent data to Chinese firms, it also has some limitations, which together with suggested remedies are discussed in an online appendix.

Our integrated dataset matching ASIE to Qin/Oriana covers 316,000 firms, while the full ASIE sample for 2001-6 contains 374,000 firms (average $T_i = 2.3$). Tables A-1 and A-2 in the online appendix contain information and descriptive statistics on the sample of firms used in our regression analysis.

3   Empirical Strategy

Our objective is to analyze the drivers behind the explosion in patent filings in China. The existing evidence is ambivalent about the factors that have contributed to the rapid rise in patent filings. On the one hand, Hu and Jefferson (2009) suggest that patenting in China is explained by increases in FDI, China's WTO accession, and improvements in the legal framework and enforceability of IP, with the latter two empirically captured by time dummies. On the other, there is a widely-held view that SIPO rubber-stamps patent filings which protect at best low-value, incremental inventions (Puga and Trefler, 2010), and that filings are largely driven by government incentives which target patenting directly (Li, 2012; Dang and Motohashi, 2015; Lei et al., 2015).

To explore the determinants of patenting in China we use the patent production function approach (Pakes and Griliches, 1980; Hall and Ziedonis, 2001) that relates a firm's patent filings to a standard set of variables, such as R&D expenditure, firm size and age. In light of the export-innovation literature, we extend the standard patent production equation to include a firm's export intensity: there is strong theoretical and empirical evidence pointing to a positive association between innovation and exporting, and if patent filings are driven by innovation we would expect exporting to predict patenting (Salomon and Shaver, 2005; Girma et al., 2008; Melitz and Trefler, 2012). Our extended specification also includes financial variables that are key determinants of corporate innovation activities (Brown et al., 2009, 2012; Guariglia and Liu, 2014). To test directly for the role of subsidies in the patent explosion, we include a

province-level indicator for the availability of government patent subsidies in the model.

Our main interest is in our ability to predict patent filings with SIPO by companies resident in China using the patent production function approach, which allows us to analyze the determinants of the Chinese patent explosion. To provide a benchmark against which to compare our results on the predictors of patent filings with SIPO, we use the same production function to predict patent filings by Chinese residents with the USPTO: since patent filings in the U.S. are subject to a different standard than filings with SIPO (for a detailed discussion see online appendix D), comparing the determinants of USPTO and SIPO patent filings by the same set of companies in China offers additional insights on the determinants of the patent explosion in China. More specifically, controlling for all standard determinants of firm-level innovation and patenting including a set of variables capturing financial constraints, if patenting with SIPO is driven by factors other than innovation, we expect in particular export intensity to predict filings only at the USPTO but not SIPO. In contrast, due to the policy drive to promote domestic patenting directly, we expect patent subsidies to predict filings only with SIPO but not the USPTO.

We test these hypotheses through a number of alternative empirical models which are all variations of the Pakes and Griliches (1980) patent production approach. We begin with the patenting decision, where we disregard the patent count and focus merely on the prevalence of patenting. We employ binary choice models to analyse two dichotomous outcomes, namely patenting with SIPO and patenting with the USPTO, in a standard random utility formulation (Greene and Hensher, 2010).

We address selection into our integrated dataset, a subsample of ASIE, as part of our analysis of the patenting decision by modelling selection and patenting jointly: in bivariate probit models for USPTO and SIPO patenting, respectively (results available on request), and then in trivariate probit models jointly estimating selection, patenting with the USPTO, and patenting with SIPO.[9] The formal representation of the trivariate probit model is

$$
\begin{pmatrix}
\mathbf{1}_{\{\text{sipo}\}_{ipt}} \\
\mathbf{1}_{\{\text{uspto}\}_{ipt}} \\
\mathbf{1}_{\{\text{ss}\}_{ipt}}
\end{pmatrix}
= \Phi
\begin{pmatrix}
\alpha^1 + \mathbf{INNOV}'_{ipt}\beta^1 + \mathbf{EX}'_{ipt}\gamma^1 + \mathbf{FIN}'_{ipt}\delta^1 + \mathbf{INCENT}'_{ipt}\eta^1 + \mathbf{X}'_{ipt}\theta^1 + d_p^1 + d_t^1 \\
\alpha^2 + \mathbf{INNOV}'_{ipt}\beta^2 + \mathbf{EX}'_{ipt}\gamma^2 + \mathbf{FIN}'_{ipt}\delta^2 + \mathbf{INCENT}'_{ipt}\eta^2 + \mathbf{X}'_{ipt}\theta^2 + d_p^2 + d_t^2 \\
\alpha^3 + \mathbf{INNOV}'_{ipt}\beta^3 + \mathbf{EX}'_{ipt}\gamma^3 + \mathbf{FIN}'_{ipt}\delta^3 + \mathbf{INCENT}'_{ipt}\eta^3 + \mathbf{X}'_{ipt}\theta^3 + d_p^3 + d_t^3
\end{pmatrix}
, \Sigma
,
$$

where $\Phi(\cdot, \Sigma)$ is a multivariate normal distribution, $\mathbf{1}_{\{.\}}$ represents binary variables ('sipo' and 'uspto' for at least one patent application with SIPO and USPTO, respectively; 'ss' is the sample selection

---

[9] Addressing selection in these nonlinear models does not require an exclusion restriction from the selection equation: identification is in principle given through functional form (Greene and Hensher, 2010).

equation), and d_p^j and d_t^j are province (see below) and time fixed effects. We enter five groups of covariates to analyse the association of patenting with firm-level innovation effort (INNOV), export behaviour (EX) and financial constraints (FIN), as well as government patenting incentives (INCENT), on top of additional control variables (X) related to firm size, age, and ownership type.[10] In an additional specification we account for unobserved heterogeneity potentially distorting our results by including provincial dummies in the trivariate probit models. The results from this exercise (available on request) are qualitatively in line with those presented here.[11]

In order to gauge the reliability of our results in the face of potential endogeneity of our regressors, we estimate instrumental variable (IV) probit models adopting first or first and second lags of all variables (except firm age, ownership, and time dummies) as instruments.[12]

A second set of regressions then analyses the number of patent applications and grants by estimating nonlinear functions which relate the patent count to firm characteristics, using the same sets of covariates as above. We treat our panel as repeated cross-sections (see Bound et al., 1984), in the spirit of previous work on China by Hu and Jefferson (2009), but like these authors consider fixed effects Poisson models for robustness (see below). In empirical practice the choice between different approaches is primarily driven by the 'overdispersion' problem of the Poisson estimator (Cameron and Trivedi, 2005; Hilbe, 2011). The Negative Binomial estimator enables us to introduce a separate dispersion parameter κ to overcome this issue:[13] the formal model representation of these estimators is

$$Pr(Y_{it} = y_{it}) = \left[exp(-\lambda_{it})\lambda_{it}^n\right]/(y_{it}!)$$

$$Pr(Y_{it} = y_{it}) = \frac{\Gamma(y_{it} + \lambda_{it}^{1-c}/\kappa)}{y_{it}!\Gamma(\lambda_{it}^{1-c}/\kappa)} \left(1 + \kappa\lambda_{it}^c\right)^{-\lambda_{it}^{1-c}/\kappa} \left(1 + \lambda_{it}^{-c}/\kappa\right)^{y_{it}}$$

with y_{it} the patent count and $\lambda_{it} = exp(\mathbf{Z}_{it}'\boldsymbol{\varphi})$, where for convenience of notation we have expressed the five sets of covariates and dummies detailed above with matrix $\mathbf{Z}$ and their respective coefficient vectors with $\boldsymbol{\varphi}$.

We also present results from a fixed effects (FE) Poisson model, where the inclusion of firm fixed effects

---

[10]Full details of all variables and controls included in the models are contained in the online appendix.

[11]It is well-known that the inclusion of a large number of fixed effects in nonlinear models creates serious bias due to the incidental parameter problem. This problem should not create any difficulties for a mere 30 province dummies, however China's vast economic heterogeneity creates a separate problem here in that nine (two) provinces have no firms with any patent applications with USPTO (SIPO) over the 4-year sample period, which means that firms from these provinces are dropped.

[12]Additional analysis (results available on request) replaces the dependent variable of at least one patent application with that of at least one granted patent, which can act as a basic proxy for the quality of innovations — results are qualitatively identical.

[13]Tests for the statistical significance of κ reject the Poisson estimator in favour of the NegBin alternative in all cases.

limits the sample to 'innovating firms.' This reduces the number of observations from 804,766 to 507 (170 firms) in the USPTO and 7,113 (2,327 firms) in the SIPO analysis, but allows us to account for time-invariant unobserved heterogeneity thus giving an interpretation closer to causality than in our pooled regressions.

In our analysis innovation effort is proxied by R&D expenditure. We employ contemporaneous R&D expenditure (Pakes and Griliches, 1980) deflated by employment[14] to avoid confounding the R&D effect with that of firm size (Hall and Ziedonis, 2001). Log R&D expenditure per worker is entered as linear and squared terms to allow firms in different tails of the distribution to impact patenting decisions and patent counts differentially.

Inspired by the export-innovation literature we include the export intensity of a firm (export deflated by sales) as an additional determinant of patenting. Our regressions separately control for firms with zero exports and for 'pure exporters' (Defever and Riano, 2013) with export intensity in excess of 90% — this cut-off is based on investigating a kernel density estimate for this variable.

Recent work by Brown et al. (2009) and Brown et al. (2012) on advanced economies and Guariglia and Liu (2014) on China highlighted the importance of financial constraints as key determinants of corporate innovation strategies. We therefore include measures for firm liquidity, leverage and cash flow as additional covariates.[15]

The municipal government of Shanghai started handing out patent filing subsidies in 1999 and by 2007, 80% of Chinese provinces had adopted such subsidy schemes (Dang and Motohashi, 2015). There are substantial differences in subsidy programmes across provinces (Li, 2012) and many cities also offer their own patent subsidies (Lei et al., 2015); some programmes offer filing or examination subsidies, others pay out a cash reward only after successful grant. Some provincial and city governments fully reimburse filing and examination fees, others only reimburse a fraction of the fees. Others even determine subsidy amounts on a case-by-case basis. We use data collected by Dang and Motohashi (2015) on the presence and strength of provincial-level incentives targeting patenting directly, where our focus is on filing subsidies. This dataset substantially extends the information on subsidy schemes used in an earlier study by Li (2012) as it differentiates subsidy schemes between those that provide full or partial reimbursement of fees. It represents the most comprehensive available dataset on patent subsidies in

---

[14]We add dummies for firms with zero R&D expenditure (87% of observations).

[15]We define liquidity as the difference between a firm's current liquid assets and liabilities, normalised by total assets; and leverage as the ratio of total liabilities to total assets. Because R&D is treated as a current expense for accounting purposes we add R&D expenses to the standard measure of net cash flow (after-tax earnings plus depreciation) to obtain gross cash flow (see Himmelberg and Petersen, 1994); this cash flow variable is then normalised by total assets.

China. Other studies on the effect of patent subsidies have used more limited data, Lei et al. (2015) for example use data for six cities in the Jiangsu Province and Boeing and Mueller (2015) only rely on a year dummy variable to capture the introduction of a subsidy programme. Further details about the data used in our analysis and the evolution of patent subsidies across provinces over time are provided in the online appendix.

Our choice of additional firm-level controls is guided by standard suggestions in the literature, namely measures for size and age (both in logs), as well as characteristics with particular relevance for China, namely ownership type and province dummies (the latter as a robustness exercise, results available on request). Firm size is measured by employment and meant to capture possible economies of scale in patent production. In an OECD country context firm age is intended to capture the experience of older firms in the management of the patent application process (Hall and Ziedonis, 2001), however in a China emerging from a planned economy, this is an additional indicator for socialist period legacy. Ownership (our designation is based on paid-in capital share in excess of 50%, following Guariglia et al., 2011) includes two types of foreign-invested enterprises (FIEs) distinguishing those from Hong Kong, Macao and Taiwan (HMT) and elsewhere (other). We further distinguish Private, State-Owned (SOEs), Collective and Other Chinese firm types. We prefer to investigate the 'direct' effect of foreign direct investment (FDI) on patenting behaviour rather than relying on proxies suggested in the literature to capture 'knowledge spillovers' from FDI (Hu and Jefferson, 2009). We add year dummies to all models which allows us to chart the changes in patenting over time. All standard errors reported are clustered at the firm-level.

4   Descriptive Evidence

Our integrated dataset enables us to produce a number of powerful insights into Chinese patenting through simple descriptive statistics. Tables 1 and 2 list the top-10 companies patenting with the USPTO and SIPO, respectively. These tables are constructed using the patent data for the entire time horizon 1985 to 2006 for the firms in our integrated dataset.

Table 1 illustrates the concentration of USPTO patents among a small number of companies: the top-10 assignees account for slightly less than 75% of USPTO patents. Interestingly, three companies, Hong-fujin (1), Fuzhun (3) and Futaihong (6), are subsidiaries of the Taiwanese-owned multinational Foxconn Technology Group, the world's largest contract manufacturer of 3C (Computer, Communication, Con-

sumer electronics) products. These three subsidiaries account for 35% of total USPTO patent filings in our matched dataset, adding in communications giant Huawei brings the tally to over 50%. As shown in the last column of Table 1, with the exception of Sinopec, Nuctech, and BYD, all top-10 USPTO patentees are in 3C industries. Table 2 shows SIPO patent filings, with the top-10 companies accounting for over half of all patents. Here the dominant player is Huawei, filing nearly a quarter of SIPO patents, whereas only one Foxconn subsidiary, Hongfujin, is among the top-10. Again, with the exception of Sinopec, BYD and Baoshan, all companies listed in Table 2 are in 3C industries. Note that there is a significant overlap of companies in Tables 1 and 2: six companies appear in both lists, with four of these in 3C industries.

Apart from asking who patents, the question of what is patented is equally important. We classify USPTO and SIPO patents according to the type of innovation they protect: product or process innovation or a combination of the two. There is a common perception in the literature that patents protecting product inventions reflect genuine innovations whereas process patents are of less innovative content as they only indicate new ways of producing some output by existing means. We read random subsamples of 1,900 USPTO and 980 SIPO patents.[16] Table A-3 in the online appendix shows a breakdown of patents filed by Chinese residents according to the innovation type they protect. For USPTO patents nearly half cover product innovations and only 20% process innovations. The pattern looks different in the case of SIPO patents: merely 30% protect product innovations and 37% process innovations. This analysis suggests that inventions that are patented in China but not in the U.S. are more likely to protect process innovations. In contrast, results for USPTO patents indicate that the share of patents protecting product innovations is substantially higher.

Although there is clear evidence for substantial concentration of patenting among a small number of firms with either jurisdiction, we can also distinguish the observable characteristics between firms which (a) do and do not patent, and in turn between those which (b) patent with SIPO and the USPTO. Table 3 provides the respective unconditional mean comparison with associated one-sided t-tests. The columns on the left compare characteristics for patentees with non-patenting firms, highlighting the correlation between patenting and innovation effort (R&D expenditure). While export intensity is qualitatively similar, non-patenting firms have a higher propensity to be non-exporters or pure exporters — both of the latter findings ring true with reference to work on productivity and exporting (Melitz and Trefler, 2012; Defever and Riano, 2013). Patenting firms are larger, older and have higher liquidity than non-

---

[16]In the case of SIPO patents claims must be retrieved from the original patent documents which are only available in Chinese.

patenting firms, while state incentives to patent are higher in provinces where patenting firms are located. These results echo the findings of Guariglia and Liu (2014) who use new product sales as an indicator of innovation. Our simple analysis of means also finds significant differences between firms patenting (a) with the USPTO or (b) (only) with SIPO (in the columns on the right of the table): among the characteristics which distinguish USPTO patentees from those firms which patent only domestically, the higher R&D expenditure, export-to-sales ratio, and firm size are particularly noteworthy. A number of characteristics are also surprisingly unimportant in this comparison, notably the financial variables (except for cash flow) and the provincial-level subsidies for patent applications.

5 Results

5.1 Patenting decision

We begin our discussion with the empirical results for the (binary) patenting decision. Table 4 reports results for the 4-year sample for which R&D expenditure is observed.[17] In all cases the data for the ASIE sample (ASIE-Qin/Oriana match and ASIE-only firms) are used and near the top of each table panel we indicate whether we account for selection into the integrated ASIE-Qin/Oriana sample.

Columns [1] and [2] represent simple probit models for the patenting decision with SIPO or USPTO, while in column [3] we add a sample selection equation for ASIE-Qin/Oriana firms which is estimated jointly with the two patenting decision equations (results for bivariate probit estimating selection and SIPO or USPTO patenting jointly yield qualitatively very similar results and are available on request). The trivariate probit results suggest that our matched-sample regression does not suffer significant selection bias and that estimating patenting equations for SIPO and USPTO separately only affects estimation and inference at the margin. The remaining columns then attempt to counter concerns over endogeneity by instrumenting with the first lag and first and second lags in columns [4]-[6] (in column [6] we additionally instrument R&D expenditure using first lags). In the absence of obvious external instruments, these specifications provide some indication of the robustness of our main findings in column [3] to endogeneity concerns. Note however that diagnostic tests yield diverging results in the SIPO and USPTO models which suggest that for the SIPO equation our instrumentation strategy violates the exclusion restriction and should therefore be interpreted with caution.

---

[17]Appendix Table F-1 shows the linear probability results.

Conditioning on pure and non-exporters, our various SIPO models suggest a significant negative relationship between export intensity and patenting behaviour, which is in stark contrast to the findings in the existing literature. There is further a significant positive relationship with government incentives to file patents and the decision to apply for a SIPO patent. These models further provide evidence for a significant positive effect of innovation effort on the patenting decision, while firm size, foreign or private ownership, and financial constraints are also significant and have the expected signs.[18] On the whole the SIPO results indicate that the patenting decision is (partly) driven by government incentives, supporting the findings of Li (2012), Dang and Motohashi (2015) and Lei et al. (2015), and further that more export-intensive firms, contrary to a Melitz-type prediction of the exporting-productivity relationship, have a lower propensity to patent than their peers exporting lower shares of their output.

Turning to the USPTO models many of our results are statistically insignificant, likely due to the limited number of patentees. Nevertheless we find a significant and strong relationship between the patenting decision and innovation effort, firms size, some measures of financial constraints as well as export intensity, respectively. The coefficients on government incentives are uniformly low and statistically insignificant. Coefficients on export intensity are positive and large but not uniformly statistically significant across all models.

We further highlight those covariates for which there is a statistically significant difference for coefficients between the SIPO and USPTO equations: most strikingly, the export-innovation nexus is positive (though not necessarily statistically significant) and thus in line with the literature for USPTO equations, while filing subsidies are now even negative (in our IV models), albeit statistically insignificant. Results for indicators of financial constraints show similar deviation between SIPO and USPTO patentees, though only in the IV specification with the smallest sample size (column [5]), which is also the specification where results for patent subsidies deviate statistically significantly. We obtain qualitatively similar results when including a set of 2-digit SIC industry dummies to confirm that despite the dominance of the ITC sector our results are not driven by sector of operation.[19]

What are the quantitative implications of the differences detected between SIPO and USPTO patentees?

---

[18]The coefficient on the cash-flow variable deviates from the existing literature on China (e.g. Guariglia and Liu, 2014) in that firms do not appear credit-constrained. Our analysis investigates patents (for SIPO: 0.39% of observations are non-zero) as opposed to (self-reported) new product sales (10.26% of observations are non-zero) in these authors' work. Hence, differences in results may be due to that fact that patented inventions commonly represent only a subset of firms' product innovations where financial constraints are potentially less relevant.

[19]We prefer the results *without* industry fixed effects since inclusion of sectoral dummies reduces the sample size in the USPTO regression by around 25%: there are no USPTO patent filings in six sectors (Leather and fur; Furniture; Paper; Printing; Rubber and Transport Equipment) which implies that there is no variation in the dependent variable for observations in these sectors and they are thus automatically dropped from the sample. There is further non-convergence in the trivariate probit model if we introduce industry dummies. Full results are available on request.

Table 5 shows the marginal effects for the coefficients shown in Table 4. For the continuous variables we focus on a hypothetical shift of a firm from the 75th to the 85th percentile of the distribution, which in the case of export intensities equates to values of 12% and 76%, respectively. The marginal effects for export intensity in the SIPO equations range between -0.1% in the probit and -0.3% in the IV probit specifications, while they are between 0.01% and 0.07% in the USPTO equations: these figures are modest in absolute terms, although we highlight the generally low propensities to patent at the top of the table. In addition, as we indicate in the columns marked 'Ratio', the export-intensity 'effect' is a multiple of the marginal effects of other firm characteristics such as firm age and size or financial constraints (note that only results for the continuous variables are directly comparable).

5.2 Patent count analysis

We now turn to the empirical analysis of patent production, which we investigate using count regression models. We present results from three different models with distinct setup and interpretation: first, we analyse a Negative Binomial for counts of patent applications with SIPO and the USPTO in columns [1] and [2] of Table 6, respectively. These were found to be favoured over standard Poisson regressions based on a direct statistical comparison (LR test). These estimates provide insights into whether firm characteristics are associated with differential numbers of patent applications between the two jurisdictions. Second, we analyse fixed effects Poisson models in columns [3] and [4], which limit the sample to 'innovating' firms with at least one SIPO or USPTO patent application over the 4-year time horizon. The interpretation of these models is whether any changes in R&D, export behaviour, financial variables, etc. *within patenting firms over time* are associated with higher or lower patent counts; since many unobserved determinants of patenting are plausibly captured by the firm fixed effects this gets us closer to a causal interpretation of the results than the previous count data models — note however that the average number of observations per firm in these FE Poisson models is merely 3.1, thus offering precious little time series variation to identify precisely any within-firm effects. Third, we move from counts of patent applications to those of *granted patents* in the analysis in columns [5] and [6]. The patent filings-to-grant-ratio for a firm can be interpreted as a first indication of the quality of its patent filings. We find that only around 63% of SIPO filings are eventually granted whereas 83% of USPTO filings are, which motivates the analysis in columns [5] and [6].

The patent count models in columns [1] and [2] show similar patterns in terms of sign and statistical significance between SIPO and USPTO patent counts as were detected in the binary choice models

of the patenting decision. Innovation effort is positively associated with higher patent counts in line with earlier findings by Hu and Jefferson (2009). Export intensity (not investigated by these authors), in contrast, indicates a clear divergence between SIPO and USPTO counts, where the export-patenting nexus is negative (positive) for the former (latter). U.S. patent count is driven by larger and younger firms in comparison with SIPO patent counts. Firm financial variables have the expected signs but differences between jurisdictions are not statistically significant. Patenting incentives have a uniformly positive impact on patent counts (not always statistically significant) — this may be counter-intuitive, but further investigation reveals that the positive USPTO coefficient is driven by firms patenting in *both* jurisdictions.[20] Firm ownership dummies indicate that all non-state-owned firm types are more prolific USPTO patentees, whereas for SIPO this is only the case among Western FIEs, consistent with earlier findings by Choi et al. (2011) — again the difference across jurisdictions is not statistically significant. All of these results are virtually identical if we use granted patent counts in columns [5] and [6] instead of application counts.

The fixed effects results provide some qualitative indications that increases in export intensity have opposite effects on SIPO and USPTO patent counts, although these results are very imprecisely estimated. Similarly the results for filing subsidies, albeit statistically insignificant in either equation, are once again in line with the previous patterns in favour of SIPO patenting.[21]

The reported coefficients, *ceteris paribus*, are differences in the logs of predicted counts for unit increases in the regressors. We also obtained incident rate ratios (IRR), which compute the relative increase or decrease (coefficients in excess of/below 1, respectively) in patent counts in response to a unit change in the regressor (reported in Table 7) — for size and age this unit change implies a doubling of the variable due to logarithmic transformation. In the models in columns [1] and [2] the *relative* IRR for export intensity yields a twelve-fold difference between SIPO (patent count reduced to 40%) and USPTO (patent count more than quadruples),[22] that for firm size an almost three-fold difference (SIPO count doubles, USPTO count quintuples). For firm age a log unit increase sees SIPO patent count drop to 85% of the previous level, and USPTO counts to 55%, a one-and-a-half-fold difference. Similar fig-

---

[20]When we limit the SIPO patent analysis to firms which do not have USPTO patents and vice versa the results for the subsidy variable are as follows: 0.903 [t=4.00] (SIPO equation), -0.307 [t=-0.88] (USPTO equation).

[21]Note that the interpretation of the firm ownership dummies is very different in these panel FE models: these estimates now indicate the impact of a *change* in ownership, and with the results driven by a small number of observations we do not report these estimates to avoid confusion.

[22]A 'unit increase' for a variable defined as a ratio between 0 and 1 is clearly difficult to interpret. For convenient interpretation we re-estimated this model using the *logarithm* of export intensity instead of the level, where a unit increase implies a doubling of the ratio. The IRR for SIPO applications is then 0.95, that for USPTO 1.64, with a (statistically significant) 1.7-fold difference between the two. The IRRs for size and age are virtually unchanged.

ures are obtained if we carry out this exercise for the models using granted patents in columns [5] and [6].[23]

6   Conclusion

What is behind the recent Chinese patent explosion? Is China transitioning rapidly from imitating technology to producing genuine innovation? What impact does the patent explosion have on the Chinese economy and on the rest of the world? While answers to these questions are of immediate concern to policy makers in China and beyond, their empirical investigation has to date been severely hampered by data limitations: there were no data available for Chinese firms that included companies' actual patent filings or that could distinguish between invention patents and the less innovative utility and design patents. We overcome these constraints and construct a dataset that contains domestic invention and U.S. utility patent filings by 316,000 manufacturing firms registered in China. We employ the data to chart the developments from 1985-2006 and to investigate the factors associated with the Chinese patent explosion over 2001-6, accounting for concerns over selection into our regression sample from survey data representative of large and medium-sized enterprises in China.

Our answer to what lies behind the Chinese patent explosion is unambiguous: a handful of companies account for the overwhelming share of patents. Does this imply there is evidence for wider technological take-off among Chinese companies? Our analysis suggests most likely not: patenting is concentrated in very few industries and even within these is undertaken by very few albeit highly active companies. What is more, the most patent-active companies both with the USPTO and SIPO operate in the ICT sector, an industry that has become notorious for its patent battles, technological standards (including standard-essential patents), and patent pools requiring firms to arm themselves with sufficiently large patent portfolios.

Our results also point to clear differences in the determinants for the patenting decision as well as patent counts between SIPO and USPTO patentees. While the latter are positively associated with export intensity as suggested by the existing literature on export behavior and innovation, we find SIPO filings to be negatively associated. This suggests that patenting with the Chinese patent office may be to a large extent driven by factors other than underlying innovative behavior: firms patenting with SIPO are found to be responding to state incentives in the form of patent subsidies. This underscores the importance of

---

[23]All magnitudes quoted are identical, with the exception of export intensity, where the difference is now seven-fold: the SIPO count reduces to 60%, the USPTO count increases by a factor of 4.5.

incentives put in place by local governments to promote patenting directly.

From a policy point of view this implies that innovation policy objectives formulated in terms of numbers of patents (such as in the recent '12th Five-Year-Plan') may not have the desired outcome: merely promoting the filing of patents that do not protect innovative technologies may create a number of unintentional adverse consequences. For instance, patent thickets — shown to exist in ICT (von Graevenitz et al., 2013) — are likely to emerge, increasing transaction costs for companies and potentially raising barriers to entry (Hall et al., 2015). Such patenting behaviour may furthermore lead to an escalation of patent litigation similar to the explosion of patent actions witnessed in the U.S. over the past decade (PWC, 2014).

Supplementary material

Supplementary material (the Appendix) is available online at the OUP website.

Funding

Acknowledgements

Mark Rogers and his important contribution to economic research on innovation and intellectual property.

References

[1] Boeing, P. and E. Mueller. 2015: Measuring patent quality in international comparison: Index development and application to China, ZEW Discussion Papers, No. 15-051.

[2] Bound, J., C. Cummins, Z. Griliches, B.H. Hall and A.B. Jaffe. 1984. 'Who Does R&D and Who Patents?' in: National Bureau of Economic Research (ed.) *R&D, Patents, and Productivity*, Ch. 2: 21-54. NBER, Cambridge, MA.

[3] Breznitz, D. and M. Murphree. 2011. *Run of the Red Queen*, Yale University Press.

[4] Brown, J.R., S.M. Fazzari, and B.C. Petersen. 2009. Financing innovation and growth: cash flow, external equity and the 1990s R&D boom. *Journal of Finance* 64: 151-85.

[5] Brown J.R., G. Martinsson, and B.C. Petersen. 2012. Do financing constraints matter for R&D? *European Economic Review* 56: 1512-29.

[6] Cameron, C. and P.K. Trivedi. 2005. *Microeconometrics: Methods and Applications*, Cambridge, UK: Cambridge University Press.

[7] Choi, S.B., S.H. Lee, and C. Williams. 2011. Ownership and firm innovation in a transition economy: Evidence from China. *Research Policy* 40: 441-452.

[8] Dang, J. and K. Motohashi. 2015. Patent Statistics: A good indicator for innovation in China? Patent Subsidy program impacts on patent quality. *China Economic Review*, 35: 137-55.

[9] Defever, F. and A. Riano. 2013. *China's Pure Exporter Subsidies*. FIW Working Paper 121. Vienna: Kompetenzzentrum 'Forschungsschwerpunkt Internationale Wirtschaft' (FIW).

[10] Fischer, W.A. and M. von Zedtwitz. 2004. 'Chinese R&D: Naissance, Renaissance or Mirage?' *R&D Management* 34: 349-65.

[11] Ganotakis, P. and J.H. Love. 2011. R&D, product innovation, and exporting: evidence from UK new technology based firms. *Oxford Economic Papers* 63: 279-306.

[12] Girma, S., H. Görg, and A. Hanley. 2008. R&D and Exporting: A Comparison of British and Irish Firms. *Review of World Economics* 144: 750-73.

[13] Greene, W.H. and D.A. Hensher 2010. *Modeling Ordered Choices*. Cambridge, UK: Cambridge University Press.

[14] Guariglia, A. and P. Liu. 2014. To what extent do financing constraints affect Chinese firms' innovation activities? *International Review of Financial Analysis*, 36: 223-40.

[15] Guariglia, A., X. Liu and L. Song. 2011. Internal Finance and Growth: Microeconometric Evidence on Chinese Firms. *Journal of Development Economics* 96: 79-94.

[16] Hall, B.H. and R.M. Ziedonis. 2001. The patent paradox revisited: an empirical study of patenting in the U.S. semiconductor industry, 1979-95. *RAND Journal of Economics* 32: 101-28.

[17] Hall, B.H., C. Helmers, and G. von Graevenitz. 2015. Technology Entry in the Presence of Patent Thickets, NBER Working Paper No. 21455. NBER, Cambridge, MA.

[18] Harris, R.I.D. and Q.C. Li. 2009. Exporting, R&D, and absorptive capacity in UK establishments. *Oxford Economic Papers* 61: 74-103.

[19] Helmers, C., M. Rogers and P. Schautschick. 2011. *Intellectual Property at the Firm-Level in the UK: The Oxford Firm-Level Intellectual Property Database.* University of Oxford, Department of Economics Discussion Paper No. 546.

[20] Hilbe, J.M. 2011. *Negative binomial regression*. 2nd edition, Cambridge, UK: Cambridge University Press.

[21] Himmelberg, C.P. and B.C. Petersen. 1994. R&D and Internal Finance: A Panel Study of Small Firms in High-Tech Industries. *The Review of Economics and Statistics* 76: 38-51.

[22] Hu, A.G. 2010. 'Propensity to patent, competition and China's foreign patenting surge.' *Research Policy* 39: 985-93.

[23] Hu, A.G. and G. Jefferson. 2009. 'A great wall of patents: What is behind China's recent patent explosion?' *Journal of Development Economics* 90: 57-68.

[24] Hu, M. and J.A. Mathews. 2008. China's national innovative capacity. *Research Policy* 37: 1465-79.

[25] Lachenmaier, S. and L. Woessmann. 2006. Does Innovation Cause Exports? Evidence from Exogenous Innovation Impulses and Obstacles Using German Micro Data. *Oxford Economic Papers* 58: 317-50.

[26] Lei, Z., Z. Sun and B. Wright. 2015. Patent Subsidy and Patent Filing in China, mimeo.

[27] Li, X. 2012. Behind the recent surge of Chinese patenting: An institutional view. *Research Policy* 41: 236-49.

[28] Mahmood, I.P. and J. Singh. 2003. Technological Dynamism in Asia. *Research Policy* 32: 1031-54.

[29] Melitz, M. and D. Trefler. 2012. Gains from Trade when Firms Matter. *Journal of Economic Perspectives* 26: 91-118.

[30] Naughton, B. 2007. *The Chinese economy: transitions and growth*, Boston, MA: MIT Press.

[31] Nordhaus, W. D. 1969. *Invention, Growth, and Welfare: A Theoretical Treatment of Technological Change*, Boston, MA: MIT Press.

[32] Pakes, A. and Z. Griliches. 1980. Patents and R&D at the firm level: A first report. *Economics Letters* 5: 377-81.

[33] Park, W. 2008. International patent protection: 1960-2005. *Research Policy* 37: 761-66.

[34] Penrose, E. 1973. International Patenting and the Less-Developed Countries, *The Economic Journal*, 83: 768-86.

[35] Puga, D. and D. Trefler. 2010. Wake up and smell the ginseng: International trade and the rise of incremental innovation in low-wage countries. *Journal of Development Economics* 91: 64-76.

[36] PricewaterhouseCoopers (PWC) 2014. 2014 Patent Litigation Study: As case volume leaps, damages continue general decline. Available at: `http://www.pwc.com/en_US/us/forensic-services/publications/assets/2014-patent-litigation-study.pdf`, last accessed 17th June 2016.

[37] Salomon, R.M. and J.M. Shaver. 2005. Learning by exporting: new insights from examining firm innovation. *Journal of Economics and Management Strategy* 14: 431-60.

[38] Subramanian, Arvind. 2011. The Inevitable Superpower: Why China's Dominance Is a Sure Thing. *Foreign Affairs* 90.

[39] Sun, Yifei. 2000. 'Spatial Distribution of Patents in China.' *Regional Studies* 34: 441-54.

[40] von Graevenitz G., S. Wagner and D. Harhoff. 2013. Incidence and Growth of Patent Thickets - The Impact of Technological Opportunities and Complexity, *Journal of Industrial Economics* 61: 521-563.

[41] Wunsch-Vincent S., M. Kashcheeva and H. Zhou. 2015. International patenting by Chinese residents: Constructing a database of Chinese foreign-oriented patent families, *China Economic Review* 36: 198-219.

[42] Yang, D. 2008. Pendency and grant rations of invention patents: A comparative study of the US and China. *Research Policy* 37: 1035-46.

[43] Zhao, M. 2006. Conducting R&D in Countries with Weak Intellectual Property Rights Protection. *Management Science* 52: 1185-99.

TABLES

Table 1: Top 10 Chinese companies filing with USPTO (1985-2006)

Rank	Company	Patents	Share	Main Industry affiliation[♮]
1	Hongfujin Precision Industry (Foxconn)	513	21.75	Electronic computer (404)
2	Huawei Technology	399	16.92	Communications equipment (401)
3	Fuzhun Precision Industry (Foxconn)	214	9.07	Electronic computer (404)
4	China Petroleum Chemical (Sinopec)	156	6.61	Petroleum, Natural Gas Exploration (079)
5	Semiconductor Manufacturing Int.	127	5.38	Electronic apparatus (405)
6	Futaihong Precision Industry (Foxconn)	100	4.24	Communications equipment (401)
7	ZTE	61	2.58	Communications equipment (401)
8	Innocom Technology (Shenzhen)	39	1.65	Communications equipment (401)
9	Lenovo	38	1.61	Electronic computer (404)
9	Nuctech	38	1.61	Special equipment (369)
10	BYD	33	1.39	Automobiles (372)
	Other	640	27.14	
	Total	2,358	100.0	

Notes: ♮ Chinese GB/T 3-digit industry code in brackets.

Table 2: Top 10 Chinese companies filing with SIPO (1985-2006)

Rank	Company	Patents	Share	Main Industry affiliation[♮]
1	Huawei Technology	15,588	23.35	Communications equipment (401)
2	ZTE	4,578	6.86	Communications equipment (401)
3	LG Electronics Appliances Tianjin	4,244	6.36	Household electrical apparatus (395)
4	Hongfujin Precision Industry (Foxconn)	3,708	5.56	Electronic computer (404)
5	China Petroleum Chemical (Sinopec)	1,977	2.95	Petroleum, Natural Gas Exploration (079)
6	AU Optronics	1,362	2.04	Electronic computer (404)
7	Lenovo	1,137	1.70	Electronic computer (404)
8	BYD	835	1.12	Automobiles (372)
9	LG Electronics Shanghai	775	1.16	CCO (409)
10	Baoshan Iron & Steel	756	1.13	Ferrous metal smelting and rolling (320)
	Other	31,781	47.77	
	Total	66,741	100.00	

Notes: ♮ Chinese GB/T 3-digit industry code in brackets. CCO – Communications, computers & other electronic equipment

Table 3: Descriptive Analysis of Patenting Behaviour

	[1] Patents					[2] Patents				
	None	Any	Diff	t-stat	p	SIPO	USPTO	Diff	t-stat	p
Innovation effort										
ln(R&D pw)	-0.037	0.218	-0.256	**-44.64**	0.000	0.212	0.306	-0.094	**-2.13**	0.017
	[0.001]	[0.011]	[0.006]	**-22.56**	0.000	[0.012]	[0.044]	[0.044]	**-2.06**	0.020
Export behaviour										
Exports/Sales	0.178	0.171	0.007	**1.68**	0.046	0.157	0.349	-0.192	**-14.02**	0.000
	[0.000]	[0.004]	[0.004]	**1.92**	0.027	[0.003]	[0.018]	[0.014]	**-10.40**	0.000
>90% Export/Sales	0.117	0.070	0.047	**12.59**	0.000	0.059	0.214	-0.154	**-13.39**	0.000
	[0.000]	[0.003]	[0.003]	**15.81**	0.000	[0.003]	[0.018]	[0.011]	**-8.43**	0.000
Zero Exports	0.724	0.536	0.188	**35.86**	0.000	0.546	0.406	0.140	**6.16**	0.000
	[0.001]	[0.006]	[0.005]	**32.22**	0.000	[0.006]	[0.022]	[0.023]	**6.23**	0.000
Size and age										
ln(Workers)	4.726	5.748	-1.022	**-79.99**	0.000	5.701	6.374	-0.673	**-11.33**	0.000
	[0.001]	[0.015]	[0.013]	**-66.68**	0.000	[0.016]	[0.058]	[0.059]	**-11.19**	0.000
ln(Firm age)	2.093	2.401	-0.307	**-29.98**	0.000	2.401	2.396	0.005	0.13	0.449
	[0.001]	[0.010]	[0.010]	**-29.90**	0.000	[0.011]	[0.043]	[0.040]	0.11	0.454
Financial constraints										
Liquidity	0.059	0.092	-0.033	**-8.88**	0.000	0.091	0.095	-0.003	-0.28	0.391
	[0.000]	[0.003]	[0.004]	**-10.29**	0.000	[0.003]	[0.012]	[0.012]	-0.27	0.395
Leverage	0.583	0.557	0.027	**7.79**	0.000	0.557	0.553	0.004	0.40	0.343
	[0.000]	[0.003]	[0.003]	**9.51**	0.000	[0.003]	[0.010]	[0.010]	0.41	0.340
Cash flow	0.105	0.099	0.006	**2.90**	0.002	0.098	0.108	-0.010	**-1.77**	0.038
	[0.000]	[0.001]	[0.002]	**3.99**	0.000	[0.001]	[0.007]	[0.006]	-1.47	0.071
Patent subsidies										
Filing	0.511	0.586	-0.075	**-15.45**	0.000	0.586	0.589	-0.002	-0.13	0.449
	[0.000]	[0.005]	[0.005]	**-15.40**	0.000	[0.005]	[0.018]	[0.019]	-0.13	0.448
Obs	797,400	7,366				6,851	515			
Firms	371,745	2,512				2,334	178			

Notes: We carry out separate two-sample t-tests in order to compare various firm-level and regional characteristics for [1] non-patenting vs patenting firms, and [2] firms patenting with SIPO vs those patenting with USPTO. The p-value indicates the probability value for a one-sided test. We test each relationship assuming equal or unequal variances across samples (though the means reported are for the former only), hence we obtain two sets of t-statistics and corresponding p values: the test statistics in the first (second) row for each variable assume equal (unequal) variances. For illustration t-statistics in bold indicate statistical significance at the 5% level.

...

Table 4: Binary Choice Models

	[1] Probit	[2] Probit	[3] Trivariate Probit	[3] Trivariate Probit	[4] IV Probit	[4] IV Probit	[5] IV Probit	[5] IV Probit	[6] IV Probit	[6] IV Probit
Dep. Variable	SIPO	USPTO	SIPO	USPTO	SIPO	USPTO	SIPO	USPTO	SIPO	USPTO
Selection			×	×						
Instruments					1st lag	1st lag	1st & 2nd lags	1st & 2nd lags	1st lag – incl. R&D	1st lag – incl. R&D
Innovation effort										
ln(R&D pw)	**0.190** [**0.010**]***	**0.208** [**0.030**]***	**0.190** [**0.009**]***	**0.206** [**0.028**]***	**0.177** [**0.020**]***	**0.201** [**0.033**]***	**0.047** [**0.014**]***	**0.198** [**0.035**]***	**0.308** [**0.059**]***	**0.350** [**0.063**]***
ln(R&D pw)^2	**0.034** [**0.009**]***	**0.055** [**0.027**]**	**0.034** [**0.008**]***	0.039 [0.025]	**0.033** [**0.009**]***	0.041 [0.030]	0.005 [0.005]	0.040 [0.029]	-0.001 [0.061]	0.150 [0.097]
Export behaviour										
Export/Sales	**-0.264** [**0.051**]***	**0.266** [**0.149**]*	**-0.268** [**0.051**]***	**0.264** [**0.138**]*	**-0.472** [**0.121**]***	0.169 [0.200]	-0.040 [0.060]	0.375 [0.276]	**-0.418** [**0.135**]***	0.242 [0.262]
>90% Export/Sales	-0.066 [0.045]	0.011 [0.104]	-0.072 [0.045]	-0.034 [0.098]	-0.011 [0.104]	-0.028 [0.176]	0.042 [0.051]	**-0.360** [**0.200**]*	-0.018 [0.111]	-0.099 [0.216]
Zero Exports	**-0.181** [**0.024**]***	0.066 [0.092]	**-0.187** [**0.024**]***	0.012 [0.086]	**-0.309** [**0.045**]***	-0.570 [0.160]	**-0.169** [**0.027**]***	0.082 [0.146]	**-0.282** [**0.046**]***	-0.008 [0.127]
Size and age										
ln(Workers)	**0.188** [**0.008**]***	**0.277** [**0.030**]***	**0.188** [**0.008**]***	**0.249** [**0.026**]***	**0.193** [**0.012**]***	**0.299** [**0.035**]***	**0.095** [**0.013**]***	**0.284** [**0.056**]***	**0.195** [**0.018**]***	**0.302** [**0.026**]***
ln(Firm age)	-0.004 [0.010]	-0.045 [0.030]	-0.005 [0.010]	-0.041 [0.028]	**-0.026** [**0.012**]**	**-0.091** [**0.038**]**	**-0.030** [**0.007**]***	**-0.098** [**0.044**]**	**-0.026** [**0.013**]*	**-0.088** [**0.033**]**
Financial constraints										
Liquidity	**0.123** [**0.037**]***	**0.234** [**0.128**]*	**0.122** [**0.036**]***	**0.215** [**0.122**]*	0.120 [0.095]	**0.498** [**0.201**]**	**-0.202** [**0.037**]***	**0.567** [**0.214**]***	0.083 [0.134]	**0.605** [**0.158**]***
Leverage	-0.043 [0.040]	0.106 [0.138]	-0.052 [0.040]	0.029 [0.131]	-0.135 [0.082]	0.305 [0.197]	**-0.226** [**0.032**]***	**0.371** [**0.210**]*	-0.156 [0.105]	**0.379** [**0.170**]**
Cash flow	**-0.247** [**0.051**]***	-0.140 [0.196]	**-0.263** [**0.051**]***	-0.272 [0.209]	**-0.271** [**0.102**]***	0.371 [0.270]	**0.095** [**0.052**]*	0.247 [0.267]	**-0.309** [**0.127**]*	0.332 [0.268]

Continued overleaf

Table 4: Binary Choice Models (continued)

	[1] Probit SIPO	[2] Probit USPTO	[3] Trivariate Probit SIPO	[3] Trivariate Probit USPTO	[4] IV Probit SIPO	[4] IV Probit USPTO	[5] IV Probit SIPO	[5] IV Probit USPTO	[6] IV Probit SIPO	[6] IV Probit USPTO
Dep. Variable	SIPO	USPTO	SIPO	USPTO	SIPO	USPTO	SIPO	USPTO	SIPO	USPTO
Selection			×	×						
Instruments					1st lag		1st & 2nd lags		1st lag – incl. R&D	
Patent subsidies										
Filing subsidy	**0.187** [0.023]***	0.089 [0.078]	**0.176** [0.021]***	0.081 [0.064]	0.572 [0.536]	0.083 [0.198]	**2.339** [0.063]***	-1.024 [0.674]	0.666 [0.884]	-0.194 [0.175]
Ownership type										
FIE (other)	**0.111** [0.040]***	0.114 [0.100]	**0.106** [0.040]***	0.119 [0.097]	0.015 [0.122]	-0.053 [0.110]	**-0.439** [0.031]***	0.212 [0.179]	-0.015 [0.192]	-0.006 [0.105]
FIE (HMT)	**0.108** [0.042]***	0.082 [0.101]	**0.100** [0.042]**	0.117 [0.095]	0.050 [0.093]	0.002 [0.105]	**-0.294** [0.029]**	0.189 [0.139]	0.033 [0.142]	0.040 [0.103]
Private	**0.059** [0.031]*	-0.118 [0.079]	**0.060** [0.031]*	-0.081 [0.074]	**0.061** [0.033]*	**-0.209** [0.079]***	**0.039** [0.017]**	**-0.147** [0.082]*	**0.056** [0.033]*	**-0.212** [0.080]***
Collective	-0.062 [0.044]	-0.236 [0.158]	-0.065 [0.045]	-0.213 [0.149]	-0.084 [0.051]	**-0.527** [0.207]**	**-0.143** [0.025]***	-0.333 [0.210]	-0.090 [0.057]	**-0.526** [0.218]**
Other	**0.132** [0.062]**	0.08 [0.153]	**0.151** [0.061]**	0.155 [0.146]	0.101 [0.067]	0.012 [0.155]	0.015 [0.036]	0.087 [0.156]	0.0921 [0.067]	0.002 [0.131]
Further Controls	×	×	×	×	×	×	×	×	×	×
Year dummies	×	×	×	×	×	×	×	×	×	×
Smith-Blundell test of exogeneity					13.14 p=.00	12.96 p=.11	53.57 p=.00	17.51 p=.03	104.8 p=.00	8.75 p=.12
Instruments Partial R^2					0.00 − 0.75	0.00 − 0.75	0.00 − 0.75	0.00 − 0.75	0.00 − 0.75	0.00 − 0.75
Hansen J Statistic							44.45 p=.00	6.45 p=.60		
Observations	804,766	804,766	804,766	804,766	637,557	637,557	459,613	459,613	637,557	637,557
Firms	374,257	374,257	374,257	374,257	300,379	300,379	240,967	240,967	300,379	300,379

Notes: The dependent variable in all models is a dummy equal to one for a firm patenting with SIPO (USPTO) in year t and zero otherwise. All models employ the 4-year dataset with observed R&D expenditure, with the top and bottom 1% of observations winsorized for R&D expenditure, firm size, export/sales ratio, firm age, as well as the three financial variables. Model [3] accounts for sample selection by ways of modelling inclusion in the integrated ASIE-Qin/Oriana sample jointly with the patenting decision (using multivariate probit). All models [4] to [6] use lagged values (either 1st lag or 1st and 2nd lag as indicated in the column header) of the three export variables, firm size, the three financial constraints variables and the patent subsidy variable as instruments. Model [6] in addition uses the first lag of the R&D expenditure variable as instrument (similarly for the squared term). All variables and 'Further Controls' are detailed in Table A-4 in the online appendix. Statistically significant coefficients and their standard errors appear in bold. We further highlight the covariates for which there is a statistically significant difference between the coefficients in the SIPO and USPTO equations. Standard errors are clustered at the firm level.

Table 5: Binary Choice Models – Marginal Effects

Variable	[1] Probit SIPO		[2] Probit USPTO		[3] Bivariate SIPO only		[3] USPTO only		[3] Both		[4] IV-Probit (1st lag) SIPO		[4] IV-Probit (1st lag) USPTO		[5] IV-Probit (1st & 2nd lag) SIPO		[5] IV-Probit (1st & 2nd lag) USPTO		[6] IV-Probit (1st lag incl. R&D) SIPO		[6] IV-Probit (1st lag incl. R&D) USPTO	
	Effect	Ratio	Effect	Ratio	Effect	Ratio	Effect	Ratio	Effect	Ratio	Effect	Ratio	Effect	Ratio	Effect	Ratio	Effect	Ratio	Effect	Ratio	Effect	Ratio
Patenting incidence																						
Unconditional	.393%		.034%		.367%		.008%		.026%		.393%		.034%		.393%		.034%		.393%		.034%	
Continuous variables																						
Export/Sales	-.093%		.011%		-.093%		.013%		.002%		-.333%		.014%		-.256%		.066%		-.286%		.014%	
ln(Workers)	.045%	-2.1	.009%	1.2	.040%	-2.3	.002%	5.2	.005%	0.5	.058%	-5.8	.024%	.6	.108%	-2.4	.025%	2.6	.060%	-4.3	.017%	3.8
ln(Firm age)	.000%	4633	.000%	73.8	.000%	-3095	.000%	212	.000%	59.2	-.001%	490	-.001%	-9.8	-.004%	61.8	-.002%	-36.7	-.001%	449	-.001%	-72.9
Liquidity	.004%	-23.2	.002%	6.7	.003%	-29.8	.001%	22.7	.001%	3.0	.005%	-62.1	.011%	1.3	.000%	-3658	.021%	3.2	.005%	-50.4	.020%	3.3
Leverage	.000%	-250	.000%	46.1	.000%	226	.000%	1817	.000%	-527	.004%	-88.8	.002%	6.0	.013%	-19.5	.006%	10.4	.003%	-76.9	.006%	11.8
Cash flow	-.009%	10.0	.000%	-38.2	-.009%	10.5	.000%	35.3	.001%	2.8	-.014%	23.2	-.006%	-2.3	-.011%	24.4	-.003%	-20.5	-.021%	12.3	-.003%	-23.4
Dummy variables																						
Zero Exports >90	-.184%	0.5	.006%	1.9	-.185%	.5	.003%	4.3	-.001%	-2.7	-.376%	.9	.008%	1.8	-.327%	.8	.003%	19.5	-.354%	.7	-.001%	-59.1
Filing subsidy	.168%	-0.6	.008%	1.4	.161%	-0.6	.001%	15.6	.006%	0.4	.570%	-0.6	-.013%	-1.0	6.906%	0.0	.008%	8.6	.436%	-0.6	-.026%	-2.5
FIE (other)	.113%	-0.8	.011%	1.0	.103%	-0.9	.004%	3.1	.010%	0.2	.047%	-7.1	.000%	-202	-.817%	0.3	.002%	36.9	.071%	-3.6	-.001%	-77.9
FIE (HMT)	.107%	-0.9	.008%	1.3	.096%	-1.0	.004%	3.5	.009%	0.3	.085%	-3.9	.006%	2.3	-.471%	0.5	.007%	9.4	.111%	-2.3	.006%	11.8
Private	.053%	-1.8	-.010%	-1.1	.057%	-1.6	-.003%	-4.2	-.003%	-0.8	.058%	-5.8	-.017%	-0.8	.033%	-7.8	-.014%	-4.8	.053%	-4.8	-.017%	-3.9
Collective	-.056%	1.7	-.015%	-0.7	-.048%	1.9	-.004%	-2.9	-.009%	-0.3	-.090%	3.7	-.033%	-0.4	-.559%	0.5	-.026%	-2.5	-.092%	2.8	-.042%	-1.6
Other	.143%	-0.6	.008%	1.5	.142%	-0.7	.005%	2.5	.012%	0.2	.135%	-2.5	.001%	12.4	.078%	-3.3	.012%	5.6	.122%	-2.1	.000%	270

Notes: We present the partial effects and related statistics for the bivariate models in Table 4. Note that in the case of the continuous variables these are *not* average partial effects for the mean or partial effects at the average, but all provide the average partial effect on the propensity to patent of moving from the 75th percentile to the 85th percentile of the distribution of the variable in question. We adopt this strategy due to the nature of the export-sales ratio variable, which is zero for the median firm. In the lower panel of the table we present the effect of a discrete change for the binary variables indicated from the base level (0). Next to the partial effects we report the ratio of the partial effect of export intensity (export-sales ratio) relative to that of each other variable; for instance, a value of -23.2 for 'Liquidity' in column [1] indicates that the partial effects of 'Liquidity' and 'Export/Sales' have opposite signs and that in terms of magnitudes the latter is 23.2 times as large as the former. We provide these ratios for the dummy variables as well, even though these are strictly not comparable.

Table 6: Count Data Models

	[1]	[2]	[3]	[4]	[5]	[6]
	NegBin		FE Poisson		NegBin	
Dep. Variable	SIPO	USPTO	SIPO	USPTO	SIPO	USPTO
Patent Applications	×	×	×	×		
Granted Patents					×	×
Innovation effort						
ln(R&D pw)	**1.122**	**1.049**	**0.127**	0.079	**1.214**	**1.079**
	[0.176]***	**[0.139]*****	**[0.056]****	[0.104]	**[0.189]*****	**[0.146]*****
ln(R&D pw)^2	**0.336**	**0.295**	**0.057**	-0.083	**0.393**	**0.376**
	[0.094]***	**[0.118]****	**[0.029]****	[0.112]	**[0.101]*****	**[0.128]*****
Export behaviour						
Exports/Sales	**-0.978**	**1.514**	-0.564	0.782	-0.498	**1.623**
	[0.429]**	**[0.566]*****	[0.653]	[1.134]	[0.378]	**[0.614]*****
>90% Export/Sales	0.119	-0.420	0.220	-0.294	0.040	-0.451
	[0.296]	[0.385]	[0.395]	[0.725]	[0.337]	[0.402]
Zero Exports	-0.103	0.582	**-0.300**	**-0.854**	0.277	0.624
	[0.334]	[0.435]	**[0.134]****	**[0.357]****	[0.238]	[0.480]
Firm size and age						
ln(Workers)	**0.643**	**1.699**	**0.214**	0.180	**0.667**	**1.715**
	[0.086]***	**[0.145]*****	**[0.109]****	[0.508]	**[0.095]*****	**[0.156]*****
ln(Firm age)	**-0.162**	**-0.566**	-0.035	**0.771**	-0.143	**-0.553**
	[0.094]*	**[0.119]*****	[0.096]	**[0.341]****	[0.100]	**[0.121]*****
Financial constraints						
Liquidity	0.536	**2.447**	0.101	-0.258	**0.884**	**2.739**
	[0.432]	**[0.690]*****	[0.202]	[0.756]	**[0.459]****	**[0.722]*****
Leverage	**0.742**	**2.364**	-0.146	0.111	**0.999**	**2.720**
	[0.443]*	**[0.622]*****	[0.358]	[0.745]	**[0.497]****	**[0.683]*****
Cash flow	**-1.181**	-0.619	0.223	0.139	**-1.407**	-0.845
	[0.541]**	[0.842]	[0.299]	[0.745]	**[0.539]*****	[0.961]
Patenting incentives						
Filing subsidy	**1.003**	**0.884**	0.863	0.239	**0.887**	**0.820**
	[0.249]***	**[0.321]*****	[0.658]	[1.023]	**[0.272]*****	**[0.337]****
Ownership type						
FIE (other)	**1.201**	**1.807**			1.066	**2.037**
	[0.722]*	**[0.476]*****			[0.779]	**[0.500]*****
FIE (HMT)	0.762	**1.378**			0.590	**1.513**
	[0.669]	**[0.434]*****			[0.701]	**[0.418]*****
Private	-0.276	**0.808**			-0.483	**1.103**
	[0.678]	**[0.341]****			[0.726]	**[0.335]*****
Collective	-0.626	**0.930**			-0.822	**1.008**
	[0.746]	**[0.462]****			[0.813]	**[0.477]****
Other	0.086	**2.319**			0.018	**2.687**
	[0.700]	**[0.618]*****			[0.760]	**[0.623]*****
Further Controls	×	×	×	×	×	×
Year dummies	×	×	×	×	×	×
LR (p-value)	0.00	0.00			n/a	n/a
$\sum |\hat{y}_i - y_i|$	0.13	0.02			0.25	0.03
AIC	51,177	4,738			37,908	4,009
LL full model	-25,559	-2,339	-7,336	-512	-18,924	-1,974
Observations	804,766	804,766	7,113	507	804,766	804,766
Non-zero obs.	0.39%	0.03%			0.28%	0.03%
Firms	374,257	374,257	2,327	170	374,257	374,257

Notes: The dependent variable in all models is the patent count with SIPO or USPTO as indicated. All variables and 'Further Controls' are detailed in Table A-4 in the online appendix. IRR reports the incidence rate ratios — see text for details. Statistically significant coefficients (10% level) and standard errors are printed in bold. In Models [3] and [4] we omit reporting coefficients for the ownership dummies since these now indicate the patent productivity of firms switching ownership, which is misleading in the general setup of our analysis. Standard errors are clustered at the firm level.

Table 7: Count Data Models: Incidence Rate Ratios (IRR)

| | [1] | [2] | | [3] | [4] | | [5] | [6] | |
| | NegBin | | | FE Poisson | | | NegBin | | |
	SIPO	USPTO	Ratio	SIPO	USPTO	Ratio	SIPO	USPTO	Ratio
IRR									
ln(R&D pw)	3.07	2.85	0.9	1.14	1.08	1.0	3.37	2.88	0.9
ln(R&D pw)^2	1.40	1.34	1.0	1.06	0.92	0.9	1.48	1.35	0.9
Export/Sales	0.38	4.55	12.1	0.57	2.19	3.8	0.61	4.55	7.5
>90% Export/Sales	1.13	0.66	0.6	1.25	0.75	0.6	1.04	0.66	0.6
Zero Exports	0.90	1.79	2.0	0.74	0.43	0.6	1.32	1.78	1.4
ln(Workers)	1.90	5.47	2.9	1.24	1.20	1.0	1.95	5.46	2.8
ln(Firm age)	0.85	0.57	0.7	0.97	2.17	2.2	0.87	0.57	0.7
Liquidity	1.71	11.55	6.8	1.11	0.77	0.7	2.42	11.54	4.8
Leverage	2.10	10.63	5.1	0.86	1.11	1.3	2.72	10.62	3.9
Cash flow	0.31	0.54	1.8	1.25	1.16	0.9	0.24	0.54	2.2
Filing subsidy	2.73	2.42	0.9	2.37	1.26	0.5	2.43	2.42	1.0
FIE (other)	3.33	6.09	1.8				2.90	6.09	2.1
FIE (HMT)	2.15	3.97	1.8				1.80	3.97	2.2
Private	0.76	2.24	3.0				0.62	2.25	3.6
Collective	0.53	2.53	4.7				0.44	2.54	5.8
Other	1.09	10.17	9.3				1.02	10.16	10.0
Obs	804,766	804,766		7,113	507		804,766	804,766	
Firms	374,257	374,257		2,327	170		374,257	374,257	

Notes: In this table we report the obtained incident rate ratios (IRRs) for the count data models in Table 6. These represent the relative increase or decrease (coefficients in excess of/below 1, respectively) in patent counts in response to a unit change in the regressor — for size and age this unit change implies a doubling of the variable due to logarithmic transformation. The columns marked ratio report the relative IRR between USPTO and SIPO equations: for instance, export intensity yields a twelve fold difference in the IRR between USPTO (patent count quadruples) and SIPO (patent count reduced to 30%). Statistical tests indicate that the IRRs between SIPO and USPTO differ for the export intensity and firm size variables in both negative binomial models of patent applications and patent grants.