

Entry and Patenting in the Software Industry

Iain M. Cockburn^{*}

Megan J. MacGarvie^{*}

October 11, 2007

We examine the effects of software patents on entry in 27 narrowly-defined classes of software products, using a dataset with comprehensive coverage of both mature public firms and small privately held firms between 1994 and 2004. Reflecting the complex economics underlying the relationship between patent protection, entry costs and industry structure, we find that patents have a mixture of effects on entry. After controlling for firm and market characteristics including market fixed effects, we find that firms are less likely to enter product classes in which there are more software patents, more potential licensors, or more ambiguous property rights. However, all else equal, firms that hold software patents related to the market are more likely to enter these markets, suggesting that the decision to invest in IP depends on the contours of the patent landscape in specific markets.

^{*}Boston University and NBER. cockburn@bu.edu, mmacgarv@bu.edu.

Support for this research and access to data were provided by LECG Inc. via an unrestricted grant from Microsoft Corporation. We thank Anne Layne-Farrar and participants in the LECG work-in-progress meetings, Shane Greenstein, Ken Simons, Sam Thompson, and Pai-ling Yin for helpful comments.

Patents are widely recognized to play an important role in determining the nature of competition in technology-intensive industries, particularly where the nature of patent rights and contracting over intellectual property is a significant determinant of economic returns to new entrants and innovators. (Gans, Hsu, and Stern (2002).) But the economic impact of patents is complex and controversial for both public policy and business practice. While patenting has increased dramatically in many industries, consistent with an acceleration in innovation and rising private value of patents, many commentators also argue that far from stimulating innovation, the surge in patenting in recent decades may in fact be harming innovation and competition.¹ Similarly, in business practice, optimism about the value of patents and the potential returns from aggressive enforcement of patent rights (see, e.g. Rivette and Klein (2000) or Davis and Harrison (2001)) has been tempered by the recognition that increased patenting also brings additional costs and uncertainty. The contours of the patent landscape are playing a critical role shaping competition and competitive strategy in a much wider range of industries than in the past, yet there is little consensus about the impact of increased patenting, and businesses face major challenges in developing and implementing effective IP strategies.

The public policy debate on patents has been loudest in industries, such as software, that are characterized by complex and cumulative innovation, and where the nature of technology and the fragmentation of patent rights pose unusually difficult challenges (see Ziedonis (2004).) Critics—many of whom are industry participants—have argued that increases in the strength of software patents over the last two decades have stifled innovation and competition by raising the costs and uncertainty associated with patent litigation, blocking development of technology,² and favoring incumbent firms with large patent portfolios over innovative new entrants. In particular, because software products are typically highly complex, potentially infringing thousands of patents, critics argue that increased patenting has created a costly “thicket” of overlapping and fragmented property rights, and that the transactions costs and contracting problems associated with this thicket are severe enough to have a substantial negative impact on innovation.

Conversely, other observers have argued that increased use of patents may lead to greater innovation and competition in software (see, for example, Smith and Mann (2004)) This may happen through obvious mechanisms such as the incentive effect of increased appropriability of returns from R&D, as well as through

more subtle mechanisms such as the role of patents as a signal of the quality of start-up firms to outside investors, greater R&D productivity as a result of increased disclosure of useful information, easier access to venture capital or other sources of finance, or more efficient transactions in knowledge in a market with explicit property rights. Mann (2005), for example, argues that patents benefit firms that are able to use them in cross-licensing negotiations.

Evidence on the “stifling” versus “stimulating” impacts of patents on innovation and competition in the software industry is, however, mixed, and difficult to interpret. Bessen and Maskin (2007) and Bessen and Hunt (2003), for example, have argued that more-and-stronger patent rights have induced a decline in R&D spending in industries affected by software patents. One of the problems with this type of study is that R&D spending may be a poor proxy for the rate of innovation. Furthermore, R&D spending is jointly determined by competitive interaction, knowledge spillovers, the nature of technological opportunities, and appropriability conditions, as well as constraints imposed by blocking IP. The interplay of these factors has proven prohibitively difficult to model and identify empirically. Cockburn (2005), for example, finds puzzlingly few patents in the economically and technologically dynamic software-intensive field of bioinformatics.

One way to avoid these difficulties is to look at the impact of quasi-natural experiments associated with legal changes to the patent system. These include the strengthening or instituting of patent systems in countries that previously had weak or non-existent formal IPRs (Sakakibara and Branstetter (2001) for Japan, Lanjouw and Cockburn (1997) for India, Moser (2005) on patent laws in the 19th century), changes in patent rights in confined to specific technologies (Scherer and Weisburst (1995) on pharmaceuticals in Italy), or a variety of other changes to patent law or patent office practice that enhance the strength of patent protection (Lerner (2002)). Hall and Ziedonis (2001) document a surge in strategic patenting in the semiconductor industry following pro-patent policy changes in the 1980s. Provided the policy changes are exogenous, these types of studies may be able to mitigate some of the endogeneity problems, but measuring their impact on firm performance remains challenging.

In this paper we focus on a direct indicator of the impact of patents: their impact on entry and industry dynamics in the software industry. Entry by new firms is closely associated with new product introductions and technological change in many industries.³ In industries, like software, with a rapid underlying pace of

technological change and an active entrepreneurial sector, any “blocking” or “stifling” effect of patents should benefit incumbents at the expense of entrants, and ought to be visible in decreased rates of entry, exit, and turnover of firms.⁴ This is also a context in which there may be “stimulating” effects on patents on rates of entry, through mechanisms such as signaling the quality of entrants’ technology in a way that improves their ability to gain access to financing (Hall (2004), Mann and Sager (2005)), or by giving entrants quick access to profits through licensing (or being acquired), should be particularly visible. After controlling for firm and market characteristics including market fixed effects, we find that firms are less likely to enter product classes in which there are more software patents, more potential licensors, or more ambiguous property rights. However, all else equal, firms that hold software patents related to the market are more likely to enter these markets. This is particularly true for firms with no experience in related markets. These findings suggest that the decision to invest in IP depends on a combination of firm characteristics and the contours of the patent landscape in specific markets.

In software, prior research has looked at a variety of indicators. Hall and MacGarvie (2006), while finding that the market value of “downstream”⁵ software producers is negatively affected relative to “upstream” players by the legal decisions expanding software patentability, show that software patents are significantly more valuable than other types of patents for the firms that hold them. In a similar vein, Noel and Schankerman (2006) quantify the cost to software firms of the potential hold-up problem associated with having to license from many rival patent holders. Lerner and Zhu (2005) find that the increased use of patents by software firms following the *Lotus v. Borland* decision was associated with improvements in firm performance (as measured, for example, by the growth of sales). Cockburn and Wagner (2007), show that internet companies filing patents were more likely to survive the collapse of the dot-com bubble after 2001 and Merges (2006) finds evidence that firms have adjusted to the presence of patents, and that effort put into acquiring patents correlates with indicators of market success. Indeed, reviewing the experience of the software industry since the late 1980s, Merges concludes that predictions that patents would “kill” the industry by choking off entry by new firms and entrenching large bureaucratic incumbents have proven to be “mostly, but not completely, wrong.” Nonetheless, concerns about the

increased cost, risk, and “stifling” impact of patents remain high, and there is little consensus as to how industry participants can most effectively respond to these challenges.

Empirical Approach

In our empirical model of entry, we follow closely the models developed by Scott Morton (1999) and Kyle (2006), and describe the probability that firm i enters market j as depending on potential profits Π_{ij} obtained from entering the market: $\Pi_{ij} = v_j + \phi_{ij}$, where v_j is variable profits and ϕ_{ij} is the fixed cost of entry. We assume that firm i will enter market j if the cost of entry is below some cutoff value F_j^* , so that $Pr(enter_{ij}) = pr(F_j^* - \phi_{ij} > 0)$

We expect the probability of entry to depend on the following factors : market-specific factors affecting F_j^* (which include demand and competition), denoted Z , firm-specific determinants of the cost of entry X , and market-specific determinants of the cost of entry M . We express the probability that firm i enters market j in year t as follows: $Pr(enter_{ijt}) = Z_{jt}\delta + X_{ijt}\beta + M_{jt}\gamma$.

In keeping with the theoretical and empirical literature on entry, we expect the probability of entering a market to be decreasing in the strength of competition and increasing in demand. Our hypotheses about cross-market variations in the costs of entry can be stated as follows: (1) entry is decreasing in the number of patents in a market, (2) entry is increasing in the number of patents held by entrants, and (3) entry is decreasing in the cost of bargaining with rival patent holders.

We do not have information on which firms are potential entrants into which markets. As a result, in the firm-level specification that is described above, we treat all firms in our sample that have not previously entered a market as potential entrants to that market. To begin, however, we present results based on data aggregated to the market level in which we examine the association between the total number entrants to a market and the extent of patenting in that market. The advantage of this specification is that it does not rely on the assumption that all firms are potential entrants. The disadvantage is that it does not allow us to compare the differential effects of

software patents on firms with different characteristics or control for firm characteristics that are correlated with entry, such as firm size and age. In what follows we therefore also estimate firm-level models of entry.

Given the reduced-form approach we take, it is important to recognize the potential biases that may be at work in our econometric estimates. In this section, we enumerate all of the major sources of bias of which we are aware, discuss the likely direction and magnitude of their impact on our estimates, and describe our empirical strategies for dealing with them. There are three types of bias with which we have developed identification strategies: 1) time-invariant market characteristics, 2) time-varying sources of bias for which we can control, and 3) time-varying sources of bias for which there are no obvious controls.

1) Time-invariant market characteristics: Markets may have persistent features that correlate with entry and patenting. We control for these factors using market fixed effects. Our estimates of the effect of patenting on entry are then derived from the “within” correlation between *changes* in patenting and *changes* in entry over time, rather than from the cross-sectional correlations at a given point in time.

2) Time-varying sources of bias for which we have controls: It is likely that there are other, time-varying sources of bias that market fixed-effects cannot eliminate. These include a) the potential for incumbents to react to the threat of entry by filing patents, b) the quality or importance of inventions in a market, and c) the stage of the product cycle.

a) *Endogenous response by incumbents to the threat of entry:* To the extent that incumbents react to the threat of entry by filing more patents, this may bias our estimates of the impact of patents on entry towards zero. That is, if markets especially prone to entry are also those in which incumbents file more patents, a substantial endogenous response by incumbents will tend to lead our reduced-form results to provide *underestimates* of the causal impact of patenting on entry.

There are in fact reasons to believe that this potential source of bias is not important. For one thing, the time it takes for a patent application to make its way through the patent office is very long relative to product development cycles in software. The average pendency period for patents relating to the markets in our sample during the time frame considered here was 2.8 years, with the market with the lowest pendency period averaging 1.4 years and the market with the longest period at 4.8 years. Software development is a very fast-moving

process, with typical development cycles measured in months rather than years. Thus, almost all of the patents in force at the time of potential entry will have been filed well in advance of any actual product launch.

Given that patent application backlogs and the grant lags that arise from them are mainly due to administrative issues at the patent office that are exogenous to the threat of entry, we use the average grant lag as an instrument for the number of patents in a market in a given year. This is described in more detail in the following section.

These relatively long patent grant lags also help us in the analysis that makes use of regime changes in patentability. These estimates make use of the fact that the expansion of software patentability took place at different times for different types of software. The evolution of patentability is discussed in more detail below, but to briefly summarize, a series of decisions in the 1990's progressively clarified the patentability of software of different types. As a result of this gradual evolution of legal and administrative doctrine over the 1980s and 90s, some categories of software patents became arguably more easily enforceable before others. For example, software used in manufacturing or otherwise tied to physical processes was covered prior to the 1996 change in guidelines, while software more generally was covered after 1996, and financial or business methods software became clearly patentable after the *State Street* decision in 1998.

This differential evolution in patentability is reflected in differences in the volume and growth rates of patenting in different patent classes within software during the 1980s and 1990s (see Figure 1 in the Appendix). Markets in which the regime change took place earlier will have seen earlier increases in the growth of patenting. Furthermore, the patents in force in the market immediately following the regime change were filed well in advance of it.⁶ To take advantage of this source of exogeneity, we estimate the effects of patents on entry in a difference-in-differences type of analysis. for three time periods: pre regime-change, immediately after the regime change, and four years after the regime change. We hypothesize that the effects will be greatest during the period immediately following the regime change.

Obviously, these trends confound changes in patentability and the perceived value of patents with technological change in that category, but we believe that cross-category variation over time in the availability and effectiveness of patents was largely driven by exogenous changes in legal doctrine, and that this is a useful

source of identification of the economic impact of software patents. We also use specific regime changes at the patent office as instruments in the IV analysis that follows.

b) *The stage of the product life cycle:* In the literature arising from Gort and Klepper (1982), progressive innovation in a technology plays a crucial role in determining the rate of entry and the number of firms in the market. This literature envisages the evolution of a market in the following way. In the initial stage, a ground-breaking innovation emerges and the innovator enjoys temporary monopoly profits that attract entry. Entrants carry out both product R&D, which stochastically results in related innovations or modifications of the new product, and process R&D, which lowers costs. Over time, as the results of the product R&D are imitated, and the process R&D reduces costs and product prices, fewer and fewer entrants are able to realize positive profits and the rate of entry falls while the rate of exit rises.

We speculate that an increase in the strength of software patents may affect entry and exit over the product life cycle in (at least) the following ways. (1) It may lengthen the first stage in which the initial innovator enjoys a monopoly. (2) It may reduce the rate of entry, as imitation becomes more difficult (for example, if blocking IP makes sequential improvements more costly). (3) By reducing uncertainty about the entrants' quality and acting as a signal to venture capitalists, patents may result in better screening of potential entrants by venture capitalists thus raising the survival rate of the firms that do enter.

A challenge of identification arises from the fact that Gort and Klepper document a surge in the rate of patenting in a market in last stage of the product cycle, when entry is low.⁷ This observation is likely to be less relevant in fast-moving technologies with short product lifecycles such as software, where patenting strategy is increasingly directed towards "upstream" and basic technology, and firms are under increasing pressure to patent "early and often." Nonetheless, if we observe a negative correlation between patents in a market and the rate of entry, it is therefore possible that we are picking up spurious correlation associated with the stage of the product cycle.

To address this issue, we control of the stage of the product cycle using the modal citation lag to patents in the product class as an indicator of the maturity of the technology (see the appendix for a list of product classes).⁸ Since the number of citations to a patent is a function of the number of potential citations, we estimate

the modal lag using a framework that adjusts for this effect. For each product class and citing-cited year pair, we compute the citation frequency, or ratio of actual to potential citations (see Jaffe and Trajtenberg (1999)), and then identify the citation lag (citing year – cited year) with the highest citation frequency for a given product class and citing year.⁹ If the modal lag in a product category is short, it implies that the most highly cited patents in that market were granted recently, which suggests that the market is at a relatively early stage of the product cycle.¹⁰ Table A.1 (in the Appendix) lists the mean modal lags by product class for 1994-2004.¹¹

c) *The technological importance of inventions in a market*: Gilbert and Newbery (1982) describe a monopolist's incentive to pre-empt entry by filing patents, including "sleeping patents" on inventions that are not commercialized. Gilbert and Newbery state that "preemption would be very hard to identify ... because it is difficult to distinguish product development that is the result of superior foresight and technological capabilities from development that is motivated by entry deterrence."¹² One of the novel features of this paper is that it takes advantage of the shifts in software patentability in the mid-nineties to identify preemptive patenting. Because software was not explicitly patentable at the beginning of our sample period, it is likely that a significant share of the observed growth in software patents represents a "filling in" of the intellectual property landscape that goes beyond the increased numbers of patent filings driven by changes in the pace of innovation or in the technological capability of firms. We present results in Table 1 which show that the entry-detering effect of patents increased dramatically after software patentability changed. If the patent-entry relationship merely reflects unobserved heterogeneity, the unobserved factors appear to have changed at exactly the same time as the legal changes which strengthened software patents.

We further attempt to disentangle pre-emptive or strategic patenting from patenting that reflects the technological capabilities of incumbents by distinguishing between the total number of patents filed by firms in a particular software market and the quality of those of patents. A market with many patents could be a market in which incumbents are very innovative, a market in which incumbents work very hard to protect intellectual property, or both. In all three cases, if entry is deterred by incumbents' behavior, then we will observe a negative correlation between patenting and entry rates. However there are very different implications of such a finding for social welfare and the pace of technological change. Purely strategic patenting that shelters incumbents from

price competition and does not reflect any underlying technical change presumably lowers welfare. At the other extreme, patents that reflect innovation by incumbents may well be associated with higher welfare if dynamic gains from innovation offset static welfare losses from higher prices.¹³

Though computing welfare gains and losses is beyond the scope of this paper, we may nonetheless be able to identify which effect dominates by controlling for average patent quality. Suppose there was no “real” underlying technological progress, but incumbents nonetheless obtained a large number of patents. We would then observe a large number of patents, lower average patent quality, and lower entry. On the other hand, if there was a high rate of innovation by incumbents, we would again see a higher level of patenting and lower entry rates, but no diminution of patent quality. By controlling for patent quality we can distinguish between markets with many patents and markets with important patents, and arguably a finding of a negative association between entry rates and patenting after controlling for patent quality is consistent with some degree of strategic patenting.

We measure the average quality of patents in a market using the mean number of citations received by those patents, which is commonly interpreted as an indicator of patent value or importance.¹⁴ There are some obvious problems with this measure. The number of citations received may reflect the legal “size” of the patent in terms of the scope of its claims and the extent to which its disclosure of the invention constitutes prior art against future applications, rather than its quality in the sense of technological significance. More seriously, the ex post number of citations received is a function of the size of the pool of potential citing patents and the patenting strategy of subsequent applicants, which are likely to be endogenous to entry into the market in question. This hampers our ability to draw “clean” inferences about patent quality from this measure, but unfortunately no other indicators are easily available.

c) Time-varying sources of bias for which we have no controls: We use Instrumental Variables to deal with any remaining biases for which there are no obvious controls. To estimate the parameters of the market-level entry model with instruments, we used a moment-based count data model in which the ratio of within-group means is used to approximate a market-level effect.¹⁵ We use three different instrumental variables to control for endogeneity of patenting. The first of these is the number of non-software¹⁶ patents granted in year t that are held by the assignees holding patents relevant to the product market. This instrument is designed to pick up the fact

that if a firm has experience navigating the patent system for technologies other than software, that firm may be more likely to patent software inventions. If a market is populated by several such firms, there will be more software patents in the market for reasons unrelated to the threat of entry.

The second instrument is the average annual grant lag for patents in a market. The grant lag is the difference between the grant year of a patent and the year in which the application was filed. An increase in the average grant lag from year to year will be associated with a slower growth rate of patents for the same number of patent applications. While the grant lag is driven to some extent by the number of patent applications filed by firms, according to the USPTO website, “The length of this delay is determined by many factors, including PTO workload, budget and manpower levels, and patent printing schedules...[for example,] The 1986 patent grant data are lower than would have normally been expected due to a lack of printing funds.”¹⁷ We believe that the exogenous resource constraints predominate. In some a somewhat different context, Mehta, Rysman and Simcoe (2007) describe a procedure which uses the difference between grant date and application date as a source of exogenous variation to solve the age-year-cohort identification problem.¹⁸

The third instrument is a qualitative variable that captures variations in the legal strength of patent protection in software over time. It is equal to zero for all markets in 1994, reflecting the fact that the USPTO did not issue new guidelines on the patentability of software until 1996, when the dummy takes on a value of 1 for most markets for the rest of the sample period. It then takes on a value of 2 after 1998, reflecting the importance of the *State Street* decision in 1998.¹⁹ We are not aware of a test for weak instruments in the non-linear GMM setting. Instead, we use the “rule of thumb” F-test statistic of 10 proposed by Staiger and Stock (1997) for instrument validity in two-stage least squares. The F-statistic on our instruments’ coefficients is 34.66 in a first stage regression in which the dependent variable is the market’s patent stock.²⁰ When the average quality of the market’s patents is the dependent variable, the F statistic is 177.37.

Empirical literature on entry

The empirical literature on entry has focused on the roles of three main factors in influencing entry: (a) demand, (b) competition, and (c) entry costs. Our focus here will be on (c): that is, whether differences across software markets in the extent of patenting are associated with differences in entry costs. Cross-industry

comparisons of entry rates have yielded several interesting findings (see Geroski (1995) for a discussion). Dunne et al. (1998) contains estimates of entry rates averaging between 41.4% and 51.8% over five-year census periods for a panel of US industries between 1963 and 1982. Within-industry variation in entry rates appears to dominate between-industry variation²¹. Not surprisingly, high entry rates are associated with high rates of innovation (Acs and Audretsch (1990), Gort and Klepper (1982), Geroski (1989)). Greenstein and Wade (1998) study entry, exit and the product cycle in the commercial mainframe computer market, and our empirical approach is closely related to this paper as well as to Scott Morton (1999), which analyzes generic entry in pharmaceuticals and Kyle (2006), which studies international entry patterns in pharmaceuticals. Specifically on software, Giarratana (2004) provides a detailed case study of entry in encryption software, including the role of patents in facilitating trade in technology. Note that unlike much of the prior literature on entry, which focuses quite generally on the determinants of market structure and the welfare effects of entry, we are more narrowly focused here on testing for any association between intellectual property rights and the pace of innovation, where entry rates are used as a proxy for the pace of innovation. This reduced form approach frees us from making strong assumptions about a very complex institutional setting but limits the conclusions that can be drawn from our estimates. Our models do not, for example, deal explicitly with strategic interaction between firms and may therefore be a poor basis for estimating a behavioral response to policy shocks.

Data

Our data is based on the CorpTech directory of technology companies which covers 19,717 software companies over the period 1990-2004.²² We have detailed information on the product classes in which these firms are active, which will form the basis of our analysis of entry into new product areas. We also know the founding date of the firm, revenues and employment for most (but not all) of the firms in the dataset, the patents held by the firm, information on corporate parents, funding sources, and a number of other variables.

One of the key CorpTech variables for the purposes of this study is the “SOF category” (or categories) assigned for each firm, which indicate the software product markets in which a firm is active at a very detailed level. This variable is self-reported and can include products under development as well as products already launched. CorpTech reports more than 290 SOF categories, however many of these are quite vaguely defined, or

appear to be defined in terms of customer segments rather than in terms of a technology—e.g. “secondary school software, dental practice management software, etc.” Furthermore mapping patents to markets is a challenging and resource-intensive task. We therefore focus our analysis on 27 of these SOF-defined markets.

These 27 markets were chosen primarily on the basis of our assessment as to whether the technology/product is reasonably distinctive, and we could define a set of keywords that could be fruitfully searched in the abstract of patent documents. Clearly there is some potential for selection bias to influence our results, however we believe that the criteria used to choose these markets are independent of entry and exit dynamics and the sample of 27 SOFs does not appear to be markedly different from the other 262 in terms of firm characteristics and entry and exit rates (see the Appendix).²³

The markets that we consider are listed in Table A.2, along with counts of numbers of incumbents and entrants. The Appendix describes the process by which we selected these categories. We identify entry in these markets from information in a dataset based on surveys conducted by CorpTech.²⁴ This dataset lists firms active in each SOF-defined market every second year from 1990 to 2004., We classify a firm as an entrant if the firm has products in a SOF category after two consecutive sample years (4 years elapsed time) of not having products in that class, or is born no more than two years before its first appearance in the dataset. Some summary statistics on entry by product class are found in Table A.2.²⁵

Patents and Entry Costs

Patents that block a would-be entrant from producing or selling its product may be a significant barrier to entry. The entrant must bear additional costs of “inventing around” such patents, pay licensing fees to the patent holder, or accept potentially severe ex post penalties.²⁶ However, constructing an indicator of the significance of patents for costs of entry to a software market, or even a simple count of the number of patents covering the technologies relevant to that market, is no simple task. The Appendix to this paper describes the resource-intensive process used to match USPTO patent classifications to the CorpTech SOF categories. In short, we use a combination of text searching and reading the manual of patent classification to identify the set of key patents associated with each market. We then estimate the relationship between the rate of entry and the number of patents in the market, controlling for time-invariant characteristics of the market using fixed effects.

Note that it may not be just the absolute *number* of patents in an area that can deter entry, but also the extent to which those patents form a “thicket.” According to Shapiro (2001), “a patent thicket is a dense web of overlapping intellectual property rights that a company must hack its way through in order to actually commercialize new technology. With cumulative innovation and multiple blocking patents, stronger patent rights can thus have the perverse effect of stifling, not encouraging, innovation.” We experiment with two alternative measures of the effects of patent thickets. One of these measures is intended to capture the number of holders of potential blocking patents (i.e. the number of potential licensors with which an entrant would have to negotiate). The other seeks to measure the concentration of patent rights in a field. Both measures are computed using patent citations.

Patent citations are references to “prior art,” or existing patented technologies, listed in the patent document.²⁷ Since citations delimit the property rights represented by a patent by describing related claims contained in other patents, citations made by a patent give an indication of the extent to which a technological area is already covered by intellectual property rights and thus foreclosed to entrants who do not obtain a license. The factors determining the total cost of obtaining licenses to allow entry are complex. All else equal, we expect that the more patents that must be licensed, the higher the total cost of entry. However, particularly in complex technologies, patents are frequently bundled or pooled or jointly licensed, the number of patents blocking the would-be entrant may not present a complete picture of entry costs. Another salient feature of “thickets” is the higher costs associated with negotiating with many parties. To the extent that there are fixed costs of conducting a negotiation, having to deal with more parties will drive up costs. To capture these effects we therefore use the number of patent-holders cited by patents in a product market (the number of cited assignees) to proxy for the number of potential licensors. We hypothesize that as the potential number of licensors with which an entrant would have to negotiate increases, the costs of entry will increase. Note however that the number of licensors may not be the only determinant of entry costs: one offsetting factor may be the concentration of ownership of intellectual property. For a given number of licensors (following Noel and Schankerman (2007) and Ziedonis (2004)), we hypothesize that entry costs will be lower when the ownership of patent rights is more concentrated, because there are fewer parties with whom to negotiate. Assuming that the share of citations received by an

assignee proxies the importance of negotiating with that assignee, we postulate that in a market which has many cited assignees but where citations go disproportionately to a small number of firms, entry costs may actually be lower than in a market with fewer assignees each of which receives a similar share of total citations. To capture this effect, we calculate the four-assignee concentration index for citations for each product market in each year.²⁸

Finally, we include a measure of the *ambiguity* of the boundaries of patents in a market. Lemley and Shapiro (2006) have discussed the problems created when there is uncertainty about the validity and scope of patent rights. This is especially problematic in software, a field in which Bessen and Meurer (2008) find that patents are more than twice as likely as other patents to have claim construction appealed to the Federal Circuit, which they take as an indicator of elevated uncertainty about the boundaries of patents in software.²⁹ To capture this feature of software patents, we compute the ratio of the number of claims allowed on a patent to the number of patents cited. Our reasoning is that patents with a very large number of claims and very few citations to prior art are whose validity and scope are likely to be particularly difficult to assess.

Descriptive Statistics

Tables A.1, A.2 and A.3 provide descriptive statistics for the data set. The average SOF-defined product market in this sample had 148.2 active firms. The average number of incumbents per market grew steadily from 74.4 in 1994 to 201.9 in 2004, while the average number of entrants varied from 7.4 to 22.0 during the sample period. Our measure of the maturity of the technology, the modal citation lag to patents in the market, varied across markets from an average of 4.33 years to 7.17 years. Figure 2 plots the “turnover” in these markets over time, showing considerable turbulence, though the “classic” patterns of industry dynamics found by Klepper and others are difficult to discern. The number of patents related to each market ranged from an average of 9.31 over the sample period to almost 4000, with an average over the whole sample of 1497.96. However, patent holding was highly concentrated with only 0.7% of firms having any patents relevant to a given market in a given year, and the average firm in the sample holding a stock of 0.08 patents in a given market. Figures 3 and 4 display the share of firms holding patents in a market and the average stock of patents held by type of firm (incumbents, entrants, and non-entrants). Interestingly, the fraction of incumbents that hold patents in a market is quite similar to the fraction of entrants that hold patents, while firms that never enter a market almost never hold patents in that

market. Our measures of bargaining costs are quite interesting. The number of cited assignees per market averages 657 with a high of 1380, and a low of 15. Clearly, the average potential entrant is very unlikely to have to obtain licenses to 1498 patents from 657 different entities—only a small fraction of the total number of patents that we have identified as being relevant to a market will be applicable to a specific product. But these figures are consistent with anecdotal evidence that in complex technologies, clearing a product for launch can entail reviewing thousands of patents.³⁰ The CR4 ratio over assignees of citations to patents in the average market was 0.30, ranging from 0.21 to as high as 0.53, confirming that patent holdings are quite highly concentrated.

Market level entry model

Table 1 contains the results of the market-level model. Given the fact that the number of entrants is a count truncated at zero and that there is unobserved heterogeneity across markets, a Poisson regression model with market fixed effects and robust standard errors is used to estimate the parameters. Wooldridge (1997) shows that, under minimal assumptions, the fixed-effects Poisson estimator is consistent and asymptotically normal.³¹ In Table 1, we present estimates from a Poisson model in which the dependent variable is the number of entrants in market j in year t , and the explanatory variables include the number of incumbent firms in the market and the number of incumbents squared, plus the growth in revenues in the market over the previous two years.³² To control for the concentration of the market, we include a proxy for the four-firm concentration ratio.³³ Because high levels of this variable can either indicate a market that is highly concentrated due to high barriers to entry or a new market ripe for entry, we include this variable in quadratic form. To control for the stage of the product cycle, we include a set of dummies for each decile of the modal citation lag of patents granted in that product market.³⁴

Our key right-hand-side variables are the log of the number of patents in the market³⁵, the log of the mean number of forward citations received by patents in the market, the average number of patents held by entrants, and the log of the mean number of forward citations received by patents held by entrants. Market and entrant patents are measured as stocks, computed from annual flows of issued patents by application date using a declining balance formula with a 15% depreciation rate and omitting patents granted more than 17 years previously.

Table 1 shows that after controlling for demand with the growth of revenues, the number of incumbents enters our model with a positive sign, and the number of incumbents squared has a negative coefficient. Both are significant at the 5% level in specifications (1) and (2). Thus, when the number of firms in the market is small, increases in the size of the market are associated with increases in entry—presumably reflecting a reduction in the market power of incumbents or a reduction of barriers to entry created by network effects. For markets with larger numbers of incumbents, however, increases in the number of incumbents reduce the probability of entry, which could reflect the fact that large numbers of incumbents indicate more mature, more crowded, and less attractive markets. A similar pattern emerges in the coefficients on the four-firm concentration ratio, with a positive coefficient on the linear term and a negative coefficient on the squared term. Our proxy for the growth rate of revenues in a product market is positively and significantly related to the rate of entry. The modal citation lag coefficients (not reported in the interest of meeting the page limit, but available upon request) display a pattern in which there is an initial increase in the rate of entry as the modal lag increases, followed by a decrease and then an increase in the coefficient when the modal lag becomes very long. These coefficients are as one would expect given the relationship between the rate of entry and the stage of the product cycle.

The log of the number of patents in a market is negatively and significantly associated with the rate of entry after controlling for the mean number of citations in the market. These effects are economically as well as statistically significant. *Ceteris paribus*, a ten percent increase in the number of patents in a market is associated with approximately a 6-7% decrease in the rate of entry (see columns 2-4 of Table 1). Consistent with the “bargaining chip” and “quality signal” hypotheses, the log of the number of patents held by entrants and the average quality of these patents are positively associated with the rate of entry. A ten percent increase in the average number of patents held by entrants is associated with approximately a 0.7 increase in the rate of entry (furthermore, the Wald test of the joint hypothesis that all the patent and patent quality coefficients are zero is rejects the null at the 1% level).³⁶ The average quality of patents in the market has a very large and strongly significant negative effect on entry, with estimates of the point elasticity around -1.2. Higher numbers of citations suggest “larger” patents, which are more difficult to invent around, and more significant innovation by incumbents, both of which will tend to deter entry. Conversely, the average quality of entrants’ patents has a

positive effect on entry: entrants with higher quality patents may find it easier to bargain their way into the market.

In column 3 we include the log of the number of cited assignees in a product market and the four-assignee concentration ratio³⁷, proxies for the costs of bargaining with rival patent holders. Holding constant other market characteristics, an increase in the number of cited assignees are associated with increases in bargaining costs, and an increase in the concentration ratio is associated with a reduction in those costs. As expected, the coefficient on the number of cited assignees is negative and statistically significant (a one percent increase in the number of cited assignees is associated with a 0.6% reduction in the rate of entry). The coefficient on the four-firm concentration ratio of citations in that market is not statistically significant. This may reflect the countervailing effects of concentration in the ownership of IP – on the one hand, a fragmented market is associated with higher costs due to the difficulty of bargaining with a larger number of potential licensors, but on the other hand, concentrated markets may be more costly to enter because of the bargaining power of assignees is greater. The results in column 4 suggest that markets in which property rights are more ambiguous (as measured by the average number of patent claims per citation) see lower rates of entry, all else equal.

Columns 5 and 6 examine the impact of changes in patentability on the entry-deterring effects of patents. As discussed previously, we created three groups of firms: those for whom software was generally patentable before 1996 (manufacturing and machine-related software), those for whom 1996 represents the regime change (most other types), and those for whom the expansion of patentability came in 1998 (financial and business-methods software). Table A.7 displays our grouping of markets by timing of patentability expansion. We then included in the regression a dummy variable equal to 1 following these regime changes and 0 before them, as well as this dummy interacted with the patent stock in the market and the average citations per patent in the market. The coefficients on the interaction terms then gives us the change in the coefficient following the regime change. We see that for entrants as a whole, the relationship between the number of patents in the market and the rate of entry is exacerbated (with a significant coefficient of -.11 on the interaction of the market's patent stock and the regime change dummy). For *de novo* entrants (which we define as firms younger than 10 years old who specialize in one aggregate SOF class – e.g. AI: artificial intelligence), the effect is stronger – the coefficient on the

interaction between market's patent stock and the dummy is -0.35. This suggests that the strengthening of IP rights in software affected entry by young, specialized firms more than established companies.

Instrumental Variables Estimates

Column 7 of Table 1 contains the results from estimating a GMM model for count data using the instruments described previously.³⁸ We use the quasi-differenced, or “multiplicative feedback model” suggested by Chamberlain (1992) and Wooldridge (1997). This model can be used for consistent estimation of the parameters in the presence of inter-temporal feedback between regressors and the dependent variable.³⁹ The results from estimating this model are found in column 7 of Table 1. We instrument for the number of patents in the market and the average quality of patents in the market.⁴⁰ The results are quite similar to those obtained in the fixed-effects Poisson model, and the test of over-identifying restrictions passes with a chi-squared statistic of 113.28 (p-value of 0.17).

Net effect of patents on entry

While the marginal estimated effects of the patent variables are interesting, the net impact in changes in the overall level of patenting on entry may be more relevant to the debate about software patents and innovation. We therefore computed estimates of the net effect on entry under a variety of scenarios. First, suppose the number of patents was increased by 10% across the board, holding constant the quality of patents in a market and the quality of patents held by entrants (and all other explanatory variables). The estimated coefficients in column (3) of Table 1 imply a net effect on entry of -4.9%, with a standard error of 1.4%, which is statistically significant at the 1% level.⁴¹ Of course, without having estimated a structural model of patenting behavior and entry decisions, these reduced form coefficients can easily be misinterpreted. For example if this increase in patents represents a pure increase in the propensity to patent rather than an increase in innovation, it is unlikely that the average quality of patented innovations would remain constant, and as a result the net decrease in entry rates implied by these regressions may be an overestimate. As an extreme example, suppose the number of patents held by both entrants and incumbents increased by 10%, but the additional patents are of no technological significance and receive no citations. Then, basing our projection on the coefficients in column 3 of Table 1, our point estimates imply an *increase* of 5.9% in the number of entrants (with a standard error of 1.7%). Here, the positive effect on

entry of a decrease in the average quality of incumbents' patents overwhelms the negative net effect of increased numbers of patents. However, a more realistic scenario might be one in which the number the new patents receive only half as many citations as the existing patents. Under this scenario, our estimates imply that an across-the-board increase of 10% in the number of patents would result in essentially no impact on the number of entrants.

It is also worth noting that our model will give different predictions if the increase in patents is asymmetric for incumbents and entrants. Since it is almost certainly easier for a small entrant with a handful of patents to double its patent stock than it is for e.g. IBM to do the same, a more realistic scenario to consider may be one in which entrants' patents increase by a larger percentage than the total number of patents in the market. As an example, suppose incumbents' patents are increased by 10% while entrants' increase by 20%. Holding patent quality constant, the net effect on the number of entrants is -4.4% with a standard error of 1.6%.

Repeating these thought experiments using the estimates obtained from the instrumental variables analysis gives similar results. Taking coefficient estimates from columns 7 of Table 1 and the associated variance-covariance matrices, we find that a 10% increase in the number of patents held by both entrants and incumbents, with no change in the mean number of citations per patent, leads to a 5.4.% reduction in the number of entrants predicted (with a standard error of 1.5%). If the additional patents are of zero quality, the estimated net effect on entry is positive but insignificant, and a 10% increase in patents with the additional patents receiving 50% fewer citations is associated with a 2.6% reduction in entry in the quasi-differenced model (with a standard error of 1.0%).

Firm level entry model

Having established these general patterns, we turn to the firm-level model to examine differences in the effects of patents on different types of entrants. We estimate a Cox proportional hazard model of entry. Results from a discrete-time hazard model are also included in column 1, and they are robust to the choice of functional form. Our dependent variable $Enter_{ijt}$ equals 1 in the year that the firm enters a market, and 0 before. Firms are dropped from the regression once they have entered a market. The dataset is therefore an unbalanced panel, with 58,037 firm-year combinations and 27 markets, for a total of 1,566,999 observations.⁴² Following Berry (1992)

and Scott Morton (1999), we begin by treating all the software firms in our sample that have not previously entered a market as potential entrants. We also experimented with more restrictive definitions of the set of potential entrants, for example by defining potential entrants as those firms that have not previously entered an “adjacent” market (the same broad SOF category, i.e. AI for artificial intelligence). The results were similar.⁴³

In addition to the explanatory variables included in the market-level regressions described above, we include the following determinants of the cost of entry: the log of the stock of relevant patents held by the firm⁴⁴ (or, alternatively, a dummy equal to 1 if the firm holds patents relevant to the market), the mean number of forward citations to the firm’s patents⁴⁵, and the measures of bargaining costs. We control for firm size using a categorical measure of the level of the firm’s revenues.⁴⁶ To control for the firm’s prior experience in related markets, we include the number of adjacent markets in which the firm is active, and a count of the number of other markets (outside the broad SOF class) previously entered by the firm. Finally, we include controls for the age of the firm, the stage of the product cycle (through the modal citation lag in the SOF), and year and market fixed effects. We cluster the standard errors by firm to account for potential correlation across observations on the same firm. The parameters in Table 2 are expressed as hazard ratios (or odds ratios in the case of the discrete-time hazard model).

The results in the first column of Table 2 show that (consistent with prior research) the probability of entering market j is increasing in the number of adjacent markets the firm has already entered. Similarly, the total number of markets in which firms are active is a significant positive predictor of entry. Although we do not report the coefficients due to space constraints, the effect of the firm’s size (revenues) on the probability of entry is concave, with an inflexion point at the \$100 million to \$250 million range.

After controlling for the average quality of patents by the mean citation count per patent, the number of patents in a market is associated with a large and significant reduction in the probability of entry. A one-unit increase in $\text{Ln}(\text{Market’s Patents})$ lowers the hazard ratio by almost one half. In elasticity terms, a 1% increase in the number of patents in a market is associated with approximately a 0.4% reduction in the hazard of entry.⁴⁷ Furthermore, if a firm holds patents in a market, the hazard of entry is approximately three times higher. However, the number of patents held does not appear to matter. The positive association between the firm’s

patents and entry persists after controlling for the average quality of the firm's patents.⁴⁸ It is also interesting to note that the coefficient on the dummy indicating that a firm has patents in a market is significantly larger than the one on the dummy indicating that the firm has other patents not related to that market (see columns 4-6). This suggests that the former variable captures something about the benefits of having intellectual property related to a specific market, and not just the phenomenon that capable firms both have more patents and are more likely to enter new markets in general. As in the results from the market-level model, our measure of the concentration of citations (the four-assignee concentration ratio) and the log of the number of cited assignees are positively and significantly associated with the probability of entry, after controlling for the number and quality of patents in a market.

When we break the results down according to whether or not the firm is active in an adjacent market (columns 5 and 6 of Table 2), the results are not dramatically different.⁴⁹ However, there are some small but interesting differences in the patent coefficients: potential entrants are more negatively affected by the number of patents in a market when they are not active in an adjacent market, and the firm's patent dummies similarly have a larger effect on entry when the firm is not active in adjacent markets. This may reflect the fact that firms already active in adjacent markets have sources of bargaining power that help counteract any deterrent effect of patents. Indeed, the coefficient on the dummy for whether the firm has patents in the target market is 3.6 for firms without experience in related markets and 1.9 for firms with experience, suggesting that holding patents is more important for firms entering a market from an unrelated market. In the rightmost column of Table 2, we present results restricted to the set of firms younger than ten years and active in only one aggregate product class (what we term the *de novo* entrants). The effect of the market's patents on entry is largest in this set of firms (with an odds ratio of 0.3), however it is not statistically significant.

We also estimated the parameters of a model of exit, where exit is identified when a firm is no longer observed in a market after two sample periods of activity in that market, or when a firm leaves the sample altogether.⁵⁰ Using the same set of predictors that we use for the entry model, we found that having patents relevant to a market reduces the odds of exiting that market by approximately 30%.⁵¹ Having additional patents further reduces the probability of exit. Increases in the number of patents in a market are associated with a

reduction in the probability of exit for the average firm, but this effect is not statistically significant. The variables intended to capture patent-related bargaining costs faced by entrants—the number of cited assignees and the concentration ratio of citations over assignees—are not significantly associated with the probability of exit.

Conclusions

Controlling for the characteristics of the firm and market, we find that software firms are less likely to enter product markets in which there are more patents. All else equal, a 1% increase in the number of patents in a market is associated with approximately a 0.5% reduction in the hazard of entry for a typical firm. This result still holds after controlling for changes in the average number of citations received by incumbents' patents (as well as the size and the growth rate of the market), which suggests that patents have an entry-deterring effect above and beyond the degree to which they reflect the technological capabilities of the firms that generate them. However, it is difficult to draw unambiguous conclusions about the overall impact of increased patenting on entry from this result alone. For example, the negative impact of patents held by incumbents (and non-competitors) on entry is offset by the apparently large benefits to entrants from holding patents: we find that, all else equal, firms holding software patents associated with a market are three times more likely to enter that market, and 30% less likely to exit a market after entry.

The net impact of increased patenting on entry is therefore quite sensitive to assumptions about accompanying changes in the quality of patents and similarities or differences in the growth rates of patenting by incumbents and entrants. The table below summarizes the implications of our estimates from the different market-level models described above for the net effect on entry, using various sets of seemingly reasonable assumptions on these factors.

Predicted change in the number of entrants associated with a ten percent increase in the number of patents	
Fixed Effects Poisson model	Quasi-Differenced GMM model
Patent holdings increase by 10%, no change in mean cites per patent	
-5.40%*** (1.54%)	-5.38%*** (1.10%)
Patents holdings increase by 10%, new patents receive zero citations	
5.86%** (2.23%)	0.4% (1.9%)
Patent holdings increase by 10%, new patents receive half as many citations as existing patents	
-0.06% (1.19%)	-2.61%** (0.99%)

As the table shows, our estimated coefficient in the market-level entry model imply a wide range of effects that an increase of 10% in the total number of software patents in the economy would, all else equal, have on the net number of entrants into the average software market over two years. Under reasonable assumptions, the effect of an across-the-board increase in patenting on entry is negative, however the magnitude of the effect depends on the extent to which these patents represent important innovations. If the patents are mostly trivial (i.e. receive very few citations), or if the percentage increase in the number of patents obtained by entrants is substantially larger than that for incumbents, an increase in patenting will have less effect.

We find that the negative relationship between patents and entry intensified following increases in the legal strength of patents, and that the effects of such an increase were largest for firms that were young and specialized. We also find that increases in an indicator of the ambiguity of the boundaries of property rights is associated with reductions in the rate of entry.

References

- Acs, Z. and D. Audretsch (1990) *Innovation and Small Firms*. Cambridge, MA: MIT Press.
- Adams, J., R. Clemmons, and P. Stephan (2006) "How Rapidly Does Science Leak Out?" *NBER Working Paper No. 11997*.
- Bessen, J. and E. Maskin (forthcoming) "Sequential Innovation, Patents, and Imitation." *RAND Journal of Economics*.
- Bessen, J. and M. Meurer (2008) *Patent Failure: How Judges, Bureaucrats and Lawyers Put Innovators at Risk*, forthcoming, Princeton University Press.
- Bessen, J. and R. Hunt (2007) "An Empirical Look at Software Patents," *Journal of Economics and Management Strategy* 16, no. 1, pp. 157-89.
- Blundell, R., R. Griffith and F. Windmeijer (2002) "Individual Effects and Dynamics in Count Data Models," *Journal of Econometrics*, Vol.108, No.1, 113-131, May 2002.
- Campbell-Kelly, M. (2003) *From Airline Reservations to Sonic the Hedgehog: A History of the Software Industry*. Cambridge, MA: MIT Press.
- Chamberlain, G. (1992) "Comment: Sequential Moment Restrictions in Panel Data." *Journal of Business and Economic Statistics*, 10, pp. 20-26.
- Cockburn, I. (2005) "State Street Meets the Human Genome Project." Chapter in R. Hahn (ed.) *Intellectual Property Rights in Frontier Industries: Bioinformatics and Software*. Washington, DC: Brookings-AEI Joint Center.
- Cockburn, I. and S. Wagner (2007) "Patents and the Survival of Internet-Related IPOs." NBER Working Paper No. 13146.
- Davis, J.L. and S. Harrison (2001) *Edison in the Boardroom: How Leading Companies Realize Value From Their Intellectual Assets*. NY: Wiley.
- Dunne, T., M. J. Roberts & L. Samuelson (1988) "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," *RAND Journal of Economics*, The RAND Corporation, vol. 19(4), pages 495-515, Winter
- Gans, J., Hsu, D. and S. Stern (2002) "When Does Start-Up Innovation Spur the Gale of Creative Destruction?" *RAND Journal of Economics*, 33(4), pp571-586.]]
- Geroski, P. (1995) "What do we know about entry?" *International Journal of Industrial Organization*, 13, pp. 421-440.
- Geroski, P. (1989) "Entry and the rate of innovation." *Economics of Innovation and New Technology* 1, pp. 203-214.
- Giarratana, M. (2004) "The Birth of a New Industry: Entry by Start-ups and the Drivers of Firm Growth. The Case of Encryption Software", *Research Policy*, 35(2): 787-806.

- Gilbert, R. and D Newbery (1982) "Preemptive Patenting and the Persistence of Monopoly." *American Economic Review*, 73(2), pp. 514-526.
- Gort, M. and S. Klepper (1982) "Time Paths in the Diffusion of Product Innovations." *Economic Journal*, 92(367), pp. 630-653.
- Graham, S. J. H. and D. C. Mowery (2003) "Intellectual Property Protection in the U. S. Software Industry." Chapter in W. M. Cohen and S. A. Merrill (eds.), *Patents in the Knowledge-Based Economy*, Washington, DC: National Academies Press.
- Greenstein, S. and J. Wade (1998) "The Product Life Cycle in the Commercial Mainframe Computer Market, 1968-1982, *RAND Journal of Economics*, 29(4), pp. 772-789.
- Jaffe, A. and M. Trajtenberg (1999) "International Knowledge Flows: Evidence from Patent Citations", *Economics of Innovation and New Technology*, Vol. 8 (1999): 105-136.
- Hall, B. (2004) "Exploring the Patent Explosion." *NBER Working Paper No. 10605*.
- Hall, B. and M. MacGarvie (2006) "The Private Value of Software Patents." *NBER WP#12195*.
- Hall, B.H. and R.H. Ziedonis (2001) "The Patent Paradox Revisited: An Empirical Study of Patenting in the US Semiconductor Industry, 1979-95," *Rand Journal of Economics*, 32(1): 101-128.
- Hausman, J., B.Hall, and Z. Griliches (1984) "Econometric Models for Count Data with an Application to the Patents-R&D Relationship." *Econometrica*, 52, pp.909-938.
- Heller and Eisenberg (1998) "Can Patents Deter Innovation? The Anti-commons in Biomedical Research" *Science*, 280(5364): 698-701.
- Klepper, S. (1996) "Entry, Exit, Growth, and Innovation over the Product Life Cycle" *American Economic Review*, 86(3), pp. 562-583.
- Kyle, M. (2006) "The Role of Firm Characteristics in Pharmaceutical Product Launches." Forthcoming in *RAND Journal of Economics*.
- Lanjouw, J. and I. Cockburn (2001) "New Pills For Poor People? Empirical Evidence After GATT." *World Development*, 29(2), pp. 265-289.
- Lanjouw, J. and M. Schankerman (2004) "Patent Quality and Research Productivity: Measuring Innovation With Multiple Indicators." *Economic Journal*, 114(495), 441-465.
- Lemley and Shapiro (2005), "Probabilistic Patents", *Journal of Economic Perspectives*, Vol. 19, p. 75, 2005
- Lessig, L. (2002) Keynote address to the *Open Source Convention* on July 24, 2002.
<http://www.oreillynet.com/pub/a/policy/2002/08/15/lessig.html?page=2> (accessed May 29, 2007)
- Lerner, Josh. (2002) "150 Years of Patent Protection." *American Economic Review*, 92(2), pp. 221-225.
- Lerner, J. and F. Zhu (2005) "Software Patent Shifts after Lotus v. Borland." *NBER Working Paper No. 11168*.

- Mann, R. J. and T. Sager (2005) "Patents, Venture Capital, and Software Start-Ups." University of Texas Law School, Law and Economics Research Paper No. 57, September 2005.
- Mehta, A., M. Rysman and T. Simcoe (2007) "Identifying the Age Profile of Patent Citations", *Working paper*.
- Merges, R. P. (2006) "Patents, Entry and Growth in the Software Industry." Working paper, UC Berkeley School of Law.
- Moser, P. (2005) "How Do Patent Laws Influence Innovation? Evidence from 19th-Century World's Fairs", *American Economic Review*, vol. 95 (4) , pp. 1215-1236
- Murray, F. and S. Stern (2007) "Do formal Intellectual Property Rights Hinder the Free Flow of scientific Knowledge? An Empirical Test of the anti-Commons Hypothesis." Forthcoming at *Journal of Economic Behavior and Organization*.
- Noel, M. and M. Schankerman (2006). "[Patent Thickets and Software Innovation: Theory and Evidence from a Panel of U.S. Firms](#)", CEPR Working Paper #5701.
- Pignalberi, G. (2005) "An Interview with Donald Knuth", *Free Software Magazine*, 2005-09-04, http://www.freesoftwaremagazine.com/articles/interview_knuth (accessed May 29, 2007).
- Prusa, T. J. and J. A. Schmitz (1991) "Are New Firms an Important Source of Innovation? Evidence from the PC Software Industry." *Economics Letters*, 35(3), pp. 339-342.
- Rivette, K.G., and D. Kline (2000) "Discovering new value in intellectual property." *Harvard Business Review* 78:54-66.
- Sakakibara, M. and L. Branstetter (2001) "Do Stronger Patents Induce More Innovation?," *RAND Journal of Economics*, Vol. 32, No. 1, pp. 77-100.
- Weisburst, S. and F.M. Scherer (1995) "Economic Effects of Strengthening Pharmaceutical Patent Protection in Italy," *International Review of Industrial Property and Copyright Law*, vol. 26, pp. 1009-1024.
- Scott Morton, F. (1999) "Entry Decisions in the Generic Drug Industry" *RAND Journal of Economics*, 30(3), pp. 421-440.
- Shapiro, C. (2001) "Navigating the Patent Thicket: Cross-licenses, Patent Pools, and Standard-Setting." In *Innovation Policy and the Economy*, Vol. 1. Adam Jaffe, Joshua Lerner, and Scott Stern, (eds.), MIT Press for the National Bureau of Economic Research.
- Smith, B. L. and S. O. Mann (2004) "Innovation and Intellectual Property Protection in the Software Industry: An Emerging Role for Patents." *University of Chicago Law Review*, 71, pp. 241-264.
- Staiger, D., and J. H. Stock (1997) "Instrumental Variables Regression with Weak Instruments." *Econometrica*, 65, p. 557-586.
- Wooldridge, J. (1997) "Distribution-Free Estimation of Some Nonlinear Panel Data Models." *Journal of Econometrics*, 90, pp. 77-91.
- Ziedonis, R. H. (2004) "Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms", *Management Science*, 50: 804 - 820.

Table 1: Market-level Entry ModelPoisson Regression with market, year and modal citation lag fixed effects; Dependent variable = Number of entrants in market j in year t

	(1)	(2)	(3)	(4)	(5) All entrants	(6) de novo	(7) MFM (i.v.) model
Incumbents (in hundreds)	0.838 (0.164)***	0.704 (0.185)***	0.717 (0.181)***	0.775 (0.183)***	0.789 (0.189)***	1.270 (0.225)***	0.703 (0.092)***
Incumbents (in hundreds) squared	-0.061 (0.018)***	-0.048 (0.020)**	-0.048 (0.020)**	-0.056 (0.020)***	-0.056 (0.020)***	-0.101 (0.021)***	-0.058 (0.009)***
Growth of revenues	0.160 (0.077)**	0.117 (0.063)*	0.117 (0.066)*	0.115 (0.067)*	0.140 (0.065)**	0.207 (0.123)*	0.164 (0.045)***
Four-firm concentration ratio (sales)	0.030 (0.014)***	0.032 (0.010)***	0.040 (0.009)***	0.032 (0.010)***	0.036 (0.010)***	0.065 (0.018)***	0.009 (0.592)
Four-firm CR squared	-0.0002 (0.0001)***	-0.0003 (0.0001)***	-0.0004 (0.0001)***	-0.0003 (0.0001)***	-0.0003 (0.0001)***	-0.0005 (0.0001)***	-0.0001 (0.0000)**
Ln(Market's patents)	-0.614 (0.177)***	-0.647 (0.172)***	-0.159 (0.225)	-0.701 (0.174)***	-0.574 (0.188)***	-0.106 (0.515)	-0.612 (0.129)***
Ln(Avg Quality of market's patents)	-1.272 (0.326)***	-1.225 (0.331)***	-1.431 (0.342)***	-1.148 (0.335)***	-1.043 (0.412)**	0.373 (1.043)	-0.695 (0.258)***
Ln(number of entrants' patents)		0.064 (0.032)**	0.088 (0.031)***	0.075 (0.030)**	0.071 (0.034)**	0.166 (0.068)**	0.031 (0.029)
Ln(Avg quality of entrants' patents)		0.045 (0.030)	0.025 (0.031)	0.041 (0.030)	0.037 (0.030)	-0.089 (0.049)*	0.067 (0.019)***
Ln(# cited assignees in market)			-0.641 (0.224)***				
CR4 of citations in market			-1.630 (1.212)				
Average claims per cite in market				-0.369 (0.188)**			
D(Regime change)					1.013 (0.668)	2.341 (1.976)	
D(Regime change)*ln(market's patents)					-0.111 (0.043)**	-0.353 (0.171)**	
D(Regime change)*ln(quality)					-0.051 (0.151)	0.052 (0.365)	
Constant	7.049 (1.615)***	7.359 (1.669)***	9.168 (1.843)***	7.956 (1.662)***	6.038 (2.054)***	-5.783 (5.575)	
Observations	162	162	162	162	162	162	162

Robust standard errors, clustered by market, in parentheses. * significant at 10%; ** significant at 5%; ***significant at 1%

Table 2: Firm-level entry model
Coefficients expressed as odds ratios.

	(1) Logit	(2)	(3)	(4) Cox PH	(5) Not active in related market	(6) Active in related market	(7) de novo entrants
# Incumbents (in 100's)	2.878*** (0.333)	2.670*** (0.304)	2.745*** (0.313)	2.629*** (0.298)	3.325*** (0.498)	1.957*** (0.286)	6.169*** (2.729)
Incumbents (100's) squared	0.923*** (0.011)	0.929*** (0.010)	0.927*** (0.010)	0.930*** (0.010)	0.914*** (0.013)	0.946*** (0.014)	0.860*** (0.035)
Four-firm CR in market	1.045*** (0.008)	1.044*** (0.008)	1.044*** (0.008)	1.044*** (0.008)	1.047*** (0.008)	1.085** (0.038)	1.040*** (0.012)
Four-firm CR squared	0.999*** (0.000)	0.999*** (0.000)	0.999*** (0.000)	0.999*** (0.000)	0.999*** (0.000)	0.999*** (0.000)	0.999*** (0.000)
Age	0.955*** (0.005)	0.955*** (0.005)	0.955*** (0.005)	0.955*** (0.005)	0.959*** (0.006)	0.951*** (0.007)	0.834*** (0.023)
Other Markets	1.276*** (0.025)	1.257*** (0.021)	1.257*** (0.021)	1.256*** (0.021)	1.383*** (0.033)	1.072*** (0.022)	
Growth of Sales in market	1.177*** (0.066)	1.175*** (0.065)	1.182*** (0.066)	1.175*** (0.064)	1.129 (0.085)	1.198** (0.096)	1.247 (0.230)
Adjacent markets	1.354*** (0.041)	1.264*** (0.033)	1.264*** (0.033)	1.252*** (0.036)		1.149*** (0.020)	1.964*** (0.128)
Ln(Market's patents)	0.545*** (0.093)	0.569*** (0.096)	0.728 (0.173)	0.580*** (0.098)	0.397*** (0.095)	0.500*** (0.118)	0.330 (0.247)
Ln(Firm's patents)	1.053 (0.120)	1.201* (0.131)	1.203* (0.131)				
Dummy for firms without patents	1.394*** (0.113)	1.157 (0.107)	1.155 (0.108)	0.998 (0.102)			
ln(Avg. quality of market's patents)	0.209*** (0.081)	0.231*** (0.086)	0.205*** (0.077)	0.243*** (0.091)	0.196*** (0.106)	0.156*** (0.077)	0.054** (0.081)
Ln(Avg. quality of firm's patents)	0.500*** (0.103)	0.436*** (0.090)	0.434*** (0.090)				
ln(Cited Assignees)			0.744 (0.161)				
Citation CR4			1.057 (0.868)				
D(firm has patents in market)				2.816*** (0.544)	3.559*** (0.688)	1.920*** (0.288)	1.399 (0.823)
D(firm has patents outside market)				1.384*** (0.145)	1.415*** (0.169)	1.571*** (0.219)	1.462 (0.444)
Observations	1558444	1558520	1558520	1558520	1370118	188402	177433

Robust standard errors, clustered by firm, in parentheses. * significant at 10%; ** significant at 5%; ***significant at 1%
All regressions also include year and market fixed effects, and a set of categorical variables measuring firm revenue

Notes

-
- ¹ In recent testimony before the U.S. House of Representatives Oversight Hearing on the Patent System (February 15, 2007), Adam Jaffe stated that, “the patent system—intended to foster and protect innovation—is generating waste and uncertainty that hinder and threaten the innovative process.”
- ² For example, Donald Knuth, an eminent computer scientist, author of *The Art of Computer Programming* and inventor of TeX, has stated that “I don’t think I would have been able to create TeX if the present [patent] climate had existed in the 1970s.” (Pignalberi (2004)).
- ³ See Graham and Mowery (2003) and Campbell-Kelly (2003) for surveys of the development of the industry, and Prusa and Schmitz (1991, 1994) on entry and new products.
- ⁴ Indeed, Acs and Audretsch (1990) and Geroski (1989) have shown that innovation is positively associated with entry rates across industries. Gort and Klepper (1982) document high entry rates at the beginning of the product life cycle after a new innovation has been developed and introduced to a market.
- ⁵ Where “downstream” refers to applications and services firms, and “upstream” refers to middleware/systems software and hardware.
- ⁶ Hall and MacGarvie (2007) find a statistically significant market reaction (measured in terms of Cumulative Abnormal Returns) to the USPTO’s issuance of new guidelines on software patentability, suggesting it was not completely foreseen by the market.
- ⁷ Gort and Klepper also show that the rate of *innovation* is “at variance with the trends in patenting”, and argue that the increased rate of patenting in the final stage of the product cycle reflects strong incentives for innovative effort arising from continued growth in the size of the market, while the actual success rate of innovative effort has declined (p. 648-650).
- ⁸ See Adams, Clemmons and Stephan (2006), who use the modal citation lag to study the rate of diffusion of scientific knowledge.
- ⁹ We compute the citation frequency as the ratio of the number of observed citations to the number of potential citations. That is, if C_{kgd} is the number of citations made to patents in SOF class k in citing year g to patents granted in SOF class k in cited year d , P_{kg} is the number of patents granted in class k in year g , and P_{kd} is the number of patents granted in class k in year d , the citation frequency is $C_{kgd}/(P_{kg} P_{kd})$.
- ¹⁰ The usefulness of this variable as an indicator of the stage of the product cycle obviously depends on the assumption that the key inventions are patented, or at least that the patented inventions
- ¹¹ Another approach is to try to identify stages of the product lifecycle through turnover rates. We experimented with this by computing SOF-year turnover rates (using the formula $[\text{entrants}-\text{exits}]/(0.5(\text{entrants}+\text{exits}))$) and estimating models on subsets of the data broken out by bottom quartile, middle-two quartiles, and top quartile of turnover rates. We did not find evidence that the effect of patents on entry changes monotonically as the rate of turnover changes. The patent coefficient was quite stable across the three groups: -1.07 (2.76), -0.51 (0.24) and -1.34 (1.07).
- ¹² Gilbert and Newbery (1982), p. 525
- ¹³ A potential countervailing effect on welfare in this particular industry could result from excess product differentiation in the presence of network effects.
- ¹⁴ See Lanjouw and Schankerman (2004).
- ¹⁵ This, the GMM counterpart to the Hausman, Hall and Griliches (1984) fixed-effects Poisson estimator, is the “mean scaling model” described in Blundell, Griffith and Windmeijer (2000), p. 5. We thank Bronwyn Hall for the use of TSP code used to estimate these models.
- ¹⁶ We begin with the Graham-Mowery definition of a software patent, that is, patents in International Patent Classes G06F, G06K, and H04L. In order to be conservative about what we treat as a software patent, we augment this definition by dropping all of the 7XX USPTO classes. See Hall and MacGarvie (2006) for a discussion of different definitions of software patents.
- ¹⁷ http://www.uspto.gov/go/taf/govt/sic/explan_sic.htm, accessed May 10, 2007.
- ¹⁸ There is some evidence that the length of a patent’s grant lag is related to the quality of the patent (Johnson and Popp (2003), Regibeau and Rockett (2007)), however it should be noted that we control for the number of citations per patent in a market.
- ¹⁹ See “*State Street*” *Decision Causes “Boom” in Software Patent Filings*, <http://library.findlaw.com/1999/Mar/1/128488.html> (accessed Aug. 30, 2006). The patent in question covered a software system for mutual fund management, and though often discussed in the context of

business method patents, this decision was viewed as a major expansion of patent protection for software in general.

²⁰ All pre-determined variables, including year and market fixed effects are included in this first stage.

²¹ Geroski (1995), p.423.

²² We define software companies as the firms listed in CorpTech as having at least one product classification beginning with “SOF”, which is CorpTech’s code for software.

²³ One area in which our sample differs, however, is in terms of the average number of patents held by firms active in the market. The average firm active in one of the sample markets has 29 patents, while the average firm in a market omitted by the sample has only 18 patents, and this difference is statistically significant. Note though that this difference arises by construction: it is difficult, if not impossible to identify patents related to many of the more vaguely defined markets. In our judgment, therefore, this subset of markets is reasonably representative of software products in general.

²⁴ We thank LECG, Inc. for providing access to these data.

²⁵ One potential pitfall arises because some firms enter CorpTech several years after their founding dates, and we thus do not observe their entry. However, only a relatively small number of these firms actually enter during the period under consideration (1994-2002). We omit SOF codes in which the number of missed entries during the period is more than one standard deviation above the mean. The average number of missed entries across the categories (calculated as the share of firms that are founded after 1990 but do not appear in the sample until more than 2 years after their founding date) amounts to 12.5% of entries, and the standard deviation is 10.08.

²⁶ Infringing valid patents can present the entrant with very substantial ex post penalties, such as damages judgments (tripled in the case of “willful infringement”) or the loss in value of assets stranded in the wake of an injunction obtained by the patent holder.

²⁷ “Prior art” is not confined to patents, indeed most forms of printed publication describing the claimed invention can constitute prior art, as can public knowledge, use, or sale of the technology.

²⁸ We also experimented with using the Herfindahl index of citations across assignees to measure concentration of patent rights, but obtained very similar results to those based on the four-assignee concentration index. These results are available from the authors upon request.

²⁹ Brian Kahin (2004) cites the hearings in 2002 on *Competition and Intellectual Property Law and Policy in the Knowledge-Based Economy*, held by the U.S. Federal Trade Commission and the Department of Justice at which Frederick J. Telecky of Texas Instruments stated that: “TI has something like 8000 patents in the United States that are active patents, and for us to know what’s in that portfolio, we think, is just a mind-boggling, budget-busting exercise to try to figure that out with any degree of accuracy at all.” Kahin notes that, “If a company with TI’s resources cannot assess what they have in-house, it is difficult to expect a small company entering a market to evaluate what claims they may be facing.”

³⁰ Based on conversations with various corporate patent counsel.

³¹ Wooldridge (2002) explains that if the underlying distribution is truly Negative Binomial, the Negative Binomial estimator is more efficient than the Poisson, but if the distributional assumption is wrong, the Poisson is still consistent as long as the conditional mean is correctly specified. He writes, “On balance, because of its robustness, the Poisson QMLE has the edge over the NegBin1 for estimating the parameters of the conditional mean.” (p. 657) In practice, we found that there was essentially no difference between results obtained using a fixed-effects Negative Binomial model and those obtained from the Poisson model. The former are available upon request.

³² Unfortunately, we do not have true market-level sales data. We create a proxy for this as follows. For firm i active in market j as well as n other markets, we compute average sales per market in market j as $SALES_{ij}/n$ (the total sales of the firm divided by the number of markets in which it is active). We then add up the average sales per market for all firms active in the market. CorpTech contains a numerical sales variable as well as a categorical variable that indicates the range in which the firm’s revenues fall. A significant portion of observations on the former are missing, and we fill in these observations with the mid-point of the range indicated by the categorical sales variable.

³³ Again, we do not have market-level sales data and use the measure of sales described above.

³⁴ Dummies for the third, fifth and seventh deciles are omitted because they are collinear with the dummies for the second and fourth deciles (respectively). The value of the modal lag at the second and third decile is 4 years, the value at the fourth and fifth deciles is 5 years, and the value at the sixth and seventh deciles is 6 years.

³⁵ This is the total number of patents relevant to the market, as defined by the concordance of patent classes to SOF classes found in the appendix. While these patents are held by a set of firms that certainly includes the incumbent firms in the market, they may also be held by firms that are not active in the market.

³⁶ We also estimated a specification including the interaction of the market's patents with entrants' patents, but the coefficient on this variable was statistically insignificant and the specification was rejected by a likelihood ratio test in favor of a specification without the interaction term.

³⁷ In the reported regression results, the number of assignees is divided by 100 to reduce the number of decimal places in the reported coefficients.

³⁸ For ease of estimation, we include the modal lag and the modal lag squared instead of a full set of dummies for all the deciles of the modal lag.

³⁹ The "mean-scaling" model (Wooldridge (1997)) was also estimated, with similar results. Blundell et al. (2000) note that a problem of the quasi-differenced estimator is that when the variables of interest are highly persistent and lagged values are used as instruments, these instruments can be quite weak. As a solution to this problem, Blundell et al. propose an estimator which incorporates pre-sample information on the dependent variable into a linear feedback model. For our purposes, however, this estimator is not useful because we do not have pre-sample information on entry. Furthermore, our sample period is relatively short, especially when using lagged values of the variables, so using the 1994 entry counts as the pre-sample information would reduce our sample period to only four years: 1998, 2000, 2002 and 2004.

⁴⁰ Due to the difficulty of getting the model to converge when including dummies for the decile of the modal citation lag, we include instead the modal citation lag and the modal citation lag squared. The coefficients on these variables are 0.116 (0.049) and -0.058 (0.008).

⁴¹ This estimate is based on the results in column 3 of Table 1. The standard errors described here are computed using the delta method (via STATA's "nlcom" command).

⁴² There are 8,555 observations in which there are missing values. Thus, the dataset on which the regressions in Table 3 are based contains 1,558,444 observations.

⁴³ As an alternative, we relax the potential entrants assumption by including presence in an adjacent market as an explanatory variable in the regression.

⁴⁴ When the log of the firm's patents is included on the right-hand-side, we also include a dummy equal to 1 if the firm has no patents.

⁴⁵ These firm-level patent and citation counts are substituted for the number of patents and average citations to patents held by entrants, which were independent variables in the market-level regressions.

⁴⁶ This is a set of dummies for each category of revenue: 0 = under \$1m; 1 = \$1m - \$2.5m; 2 = \$2.5m - \$5m; 3 = \$5m - \$10m; 4 = \$10m - \$25m; 5 = \$25m - \$50m; 6 = \$50m - \$100m; 7 = \$100m - \$250m; 8 = \$250m - \$500m; 9 = Over \$500m

⁴⁷ This can be seen by recognizing that the Cox model for the hazard is $h(t) = h_0(t) \exp(x'\beta)$. If we take natural logs of both sides, and take the derivative with respect to a particular X variable X_k , we have $\delta \ln(h_i(t)) / \delta X_k = \beta$. If X enters in logs, we can interpret the coefficient on the log of X as the percentage change in the hazard associated with a 1% change in X. Since the table reports hazard ratios, coefficients can be obtained by taking the log of the hazard ratios reported in the table.

⁴⁸ However, somewhat surprisingly, there is no significant difference in the effect of the number of patents in the market on firms that do or do not have patents, as evidenced by the coefficient on the interaction between the patent dummy and the market's patent count

⁴⁹ A potential entrant is active in an adjacent market for 12% of the firm-market-year observations.

⁵⁰ We only have data on firms that exit the sample after 1999, so we restricted the analysis to the 2000-2004 period. We do not report the results here due to the inconsistency with the sample period for the entry model, however these results are available upon request.

⁵¹ The magnitude of this effect is consistent with Cockburn and Wagner's finding for survival of dot-com's.