MANAGEMENT SCIENCE
Vol. 59, No. 1, January 2013, pp. 102–121
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.1120.1618

Appropriability Mechanisms and the Platform Partnership Decision: Evidence from Enterprise Software

Peng Huang
Robert H. Smith School of Business, University of Maryland, College Park, Maryland 20742, huang@umd.edu

Marco Ceccagnoli, Chris Forman, D. J. Wu
Scheller College of Business, Georgia Institute of Technology, Atlanta, Georgia 30308
{marco.ceccagnoli@scheller.gatech.edu, chris.forman@scheller.gatech.edu, dj.wu@scheller.gatech.edu}

We examine whether ownership of intellectual property rights (IPR) or downstream capabilities is effective in encouraging entry into markets complementary to a proprietary platform by preventing the platform owner from expropriating rents from start-ups. We study this question in the context of the software industry, an environment where evidence of the efficacy of IPR as a mechanism to appropriate the returns from innovation has been mixed. Entry, in our context, is measured by an independent software vendor's (ISV's) decision to become certified by a platform owner and produce applications compatible with the platform. We find that ISVs with a greater stock of formal IPR (such as patents and copyrights), and those with stronger downstream capabilities (as measured by trademarks and consulting services) are more likely to join the platform, suggesting that these mechanisms are effective in protecting ISVs from the threat of expropriation. We also find that the effects of IPR on the likelihood of partnership are greater when an ISV has weak downstream capabilities or when the threat of imitation is greater, such as when the markets served by the ISV are growing quickly.

Key words: platform; partnership; intellectual property rights; downstream capabilities; software industry
History: Received July 22, 2009; accepted June 22, 2012, by Gérard P. Cachon, information systems. Published online in *Articles in Advance* October 24, 2012.

1. Introduction

Platform management is a central concern to many large firms in computing (e.g., Gawer 2009, Gawer and Cusumano 2002). A challenge in platform management is that the platform owner has incentives to expropriate rents from other parties who contribute to the platform (Parker and Van Alstyne 2012). However, this reduces the latter's incentives to innovate and produce for the platform, resulting in both losses for the platform owner and a decline in social welfare.[1] A number of public and private interventions have been proposed to address this problem. For example, the platform owner could cultivate a reputation for enabling complements and commit to not expropriating rents from independent suppliers (Gawer and Henderson 2007), or give up control of the platform standard (in the sense of Katz and Shapiro 1986). Although each of these commitment mechanisms has its advantages, they are difficult to operationalize in practice, and the empirical evidence of their efficacy is inconclusive.[2]

The objective of this paper is to assess the efficacy of a different approach to this issue: the use of intellectual property rights (IPR) or ownership of strong downstream capabilities to protect the independent supplier against the threat of expropriation. We study this question in the context of the software industry, an environment where platforms are pervasive and complementary innovation by independent suppliers is often critical for platform success (e.g., Evans et al. 2006). However, although these mechanisms are particularly salient to our setting, some evidence suggests that formal IPR may not be an effective means for software start-ups to appropriate the returns from their innovations and that a majority of such firms hold no patents at all (Graham et al. 2010).

[1] See Farrell and Katz (2000) for a formal analysis. These ideas have appeared in a range of papers, including Becchetti and Paganetto (2001), Heeb (2003), Nahm (2004), and Miller (2008).

[2] For example, studying the market for handheld devices, Boudreau (2010) demonstrated that giving up control of the platform standard (by sharing intellectual property and equity ownership with partners) led to little incremental innovation beyond that achieved when the platform owner simply provided outsiders with access to the platform.

We study the decisions of independent software vendors (ISVs) to enter markets complementary to an enterprise software platform, SAP. Entry, in our context, is measured by an ISV's decision to become certified by SAP and become a member of its platform ecosystem, a potentially risky process that may entail unintended information disclosure. We interpret the entry decision as reflecting the tradeoff between the expectation of higher profits associated with access to the platform's installed base and the potential risks of expropriation by the platform owner. In this context, we find evidence that IPRs and downstream capabilities are effective at protecting ISVs from the threat of expropriation. In particular, ISVs with a greater stock of formal IPR mechanisms, such as patents and copyrights, and those with stronger downstream capabilities, as proxied by trademarks and software consulting services, are both more inclined to join the platform and to do so earlier. In our baseline specification, ISVs with high levels of formal IPR are associated with a 99.8% increase in the hazard of joining the platform; those with higher levels of trademarks are associated with a 70.1% increase in the hazard of joining. Interestingly, the two appropriability mechanisms serve as substitutes to each other and the presence of one weakens the marginal effect of the other on the likelihood and timing of partnering.

We next highlight conditions under which appropriability mechanisms are likely to be particularly salient to an ISV's entry decision. As has been noted in models of platform behavior (Miller 2008) but to our knowledge not empirically tested, the likelihood that a platform owner will expropriate rents from providers of complementary products is greatest in rapidly growing markets. Thus, we expect the value of these mechanisms to platform partners to be higher in these markets. We identify how market growth conditions the value of IPR by interacting our appropriability measures with sales growth and the rate of new entry into the market, and find that the value of IPR are greatest in such markets.

A key concern with our findings is that IPR may reflect the innovativeness of an ISV rather than the effectiveness of its appropriability mechanisms. Ownership of IPR could therefore be correlated with other dimensions of unobserved firm quality—potentially biasing our results. We perform a series of tests to provide additional evidence for our interpretation of the results. We first add a control that directly measures firm innovativeness: the number of new or improved product introductions by the ISV. This is in addition to other variables included in our baseline results controlling for firm quality, such as an ISV's publications in academic journals or conferences. Second, we control for unobserved time-invariant factors that may influence the likelihood of partnership by employing panel data models that exploit within-firm longitudinal variation to identify the effects of ownership of IPR on the probability of joining a platform at a particular time. Third, we instrument for ISVs' ownership of formal IPR using an additional source of variation in our data: legal decisions that led to changes in the IP regime for software-related inventions during and immediately preceding our sample period. Our estimates remain consistent across all of these robustness analyses. Last, we note that if our IPR findings reflect unobserved ISV innovativeness, then they must do so in a particular way: namely, they do so only for ISVs that are active in rapidly growing markets. Although this alternative interpretation is possible, it is harder to identify why unobserved quality would affect entry decisions only in such markets.

2. Hypotheses Development

In our setting, an ISV has the choice to enter into an application market that is complementary to a base system that is produced by a monopolist. In keeping with many models in this literature, we view this base system as comprising a platform, which allows for indirect network effects that arise between application developers and users. We study an environment in which the application developer already has a product and a set of its own customers, and the entry decision is essentially one to produce an application for (or to join) the platform at a particular time. Thus, the decision to join the platform becomes one to adapt the ISV's existing software to make it more valuable for users of the platform; these adaptations usually involve ensuring interoperability with the platform.

Joining the platform may increase demand for the ISV's product among users of the platform. However, this decision carries with it an increased risk that the platform owner may enter into the application market itself. This risk may increase for several reasons. First and foremost, to produce for the platform and signal compatibility to users of the base product, the ISV may have to disclose product design information to the platform owner. Acquisition of this information may make it easier for the platform owner to replicate or invent around key features of the ISV's product, lowering its costs of offering competing products. Furthermore, successful entry into the complementary application market by the ISV may provide a signal of demand in the complementary market; if this demand is strong enough then it will increase the likelihood that the platform owner will enter into the complementary market.

Platform owners who enter in this way may have ex post incentives to expropriate application developers (Farrell and Katz 2000). For example, the platform owner may enter the complementary market and produce competing products, pricing or investing in the product more aggressively than a profit-maximizing independent supplier. Or it could simply demand a low price from an independent supplier as a condition for granting access to the platform. Such actions are examples of the "ex post squeeze" defined by Farrell and Katz (2000): the platform owner induces independent suppliers to offer as much surplus as possible to buyers in the complementary market so that the platform owner can extract the surplus. These actions could lower ISV profits both from the market that is complementary to the platform as well as from the ISV's standalone customers (i.e., those who are not tied to the platform). Based on these expectations of the platform owner's ex post incentives to squeeze profits from application developers, ISVs may decide not to enter the platform.

Formal appropriability mechanisms such as IPR may be one means of deterring entry by the platform owner. Patents have been highlighted in the literature as a mechanism to protect returns from proprietary knowledge. However, because of continuing legal battles about the quality of software patents and the patentability of software, there remains considerable uncertainty about the efficacy of such formal IPR in this context, particularly for start-ups.[3] Historically, copyrights have been commonly regarded as another, sometimes more effective, form of legal protection for computer software (Graham and Mowery 2003, Graham et al. 2010). However, a series of legal decisions throughout the 1980s and 1990s recently strengthened the IP protection afforded by patents while at the same time weakening that provided by copyrights (Cockburn and MacGarvie 2011, Graham and Mowery 2003, Lerner and Zhu 2007). This has led to a decline in the use of copyright as an appropriability mechanism in software (Graham and Mowery 2003).[4]

Downstream capabilities have been emphasized in the strategy and innovation literature as an alternative to facilitate the appropriation of innovation rents (Arora and Ceccagnoli 2006, Ceccagnoli and Rothaermel 2008, Cohen et al. 2000). For example, Teece (1986) suggested that when an innovation is easily imitated or invented around, profits from an innovation may be appropriated by the owners of certain manufacturing, marketing, or other capabilities required to commercialize an innovation.

Strong appropriability mechanisms increase the cost to the platform owner of entering the complementary market. For example, expectations of legal infringements may deter imitation by the platform owner ex ante. Similarly, downstream capabilities are hard to acquire through the market on competitive terms and may therefore be rare and difficult to imitate (Teece 1986). They will also reduce the potential losses in the ISV's markets in case of an ex post squeeze. As a result, ISVs that are better protected by IPR or downstream capabilities will be more likely to partner with the platform owner.

Hypothesis 1. *The stronger an ISV's mechanisms to appropriate the returns from its innovations—such as IP protection or downstream capabilities—the more likely it will partner with the software platform owner.*

We examine the possibility that the marginal returns to IPR are decreasing in the presence of downstream complementary capabilities. The intuition underlying this prediction is straightforward. As noted in Hypothesis 1, firms with a stronger appropriation strategy, of which IP protection and complementary capabilities are two key components, are more likely to join a platform. Because ISVs with strong IP protection will have lower losses in case of entry by the platform owner, the marginal value of downstream capabilities as an additional protection mechanism may be reduced.

Hypothesis 2. *The impact of an ISV's IP protection on the likelihood that it will partner with the software platform owner is lower when the ISV has strong downstream capabilities.*

The platform owner's incentives to enter into the ISV's markets depend on the fixed cost of entry and its expected payoffs. When the ISV's target markets present higher growth opportunities, incentives to enter are higher, leading to an increase in the likelihood of entry (Miller 2008). As a result, we expect that the IPR mechanisms employed by the ISV to deter entry to be more valuable—at the margin—in growing markets, and therefore to have a greater

[3] In particular, it is often argued that the novelty and nonobviousness thresholds for granting software patents tend to be very low (Hall and MacGarvie 2010). As a consequence, software patents may be challenged and found invalid in court, and thus provide a weak safeguard against imitation. This seems to be true also in the enterprise software industry, in which secrecy is often considered a far better alternative than other appropriability mechanisms (Bader 2006).

[4] A series of court decisions throughout the early to mid-1990s widened the range of patentable software inventions. Eventually, this culminated in 1996 in the Commissioner of Patents issuing guidelines for the patenting of software that allowed inventors to patent any software embodied in physical media (Hall and MacGarvie 2010). Furthermore, in 1998 and 1999 the *State Street Bank and Trust v. Signature Financial Corporation* and *AT&T v. Excel Communications* cases strengthened business method and financial patents (e.g., Hall 2009, Lerner 2002). Over the same period, a series of cases, including several copyright infringement cases brought by Lotus Development (Lerner and Zhu 2007), weakened the protection offered by copyrights. As a result, the number of granted software patents has increased dramatically, whereas the use of registered copyrights as an appropriability mechanism in software has declined (Bessen and Hunt 2007, Graham and Mowery 2003).

impact on the likelihood of partnership between the platform owner and the ISV.[5] Thus, we propose the following:

Hypothesis 3. *The impact of an ISV's IP protection on the likelihood that it will partner with the software platform owner is greater when the markets served by the ISV have higher growth.*

3. Research Setting

The context we use to test our hypotheses is the enterprise software industry. Enterprise software consolidates the diverse information needs of an enterprise's departments together into a single, integrated software program that operates on a shared database (Hitt et al. 2002). SAP AG, the largest enterprise software vendor by revenue (SAP 2009), provides a suite of products and a set of application programming interfaces to facilitate third party integration.

We stress distinct features of our research context that naturally set our study within the platform framework. First, the value chain of SAP and its related applications have a one-to-many structure: there are literally hundreds of ISVs that produce products that are certified for use with SAP over our sample period. Second, there is a very large installed base for SAP software and complementary applications—recent estimates suggest a number over 41,000 (Pang 2007). The size of this user base suggests a benefit to users and application developers for coordinating on this platform: users who make the technical and organizational investments in the SAP platform can spread these sunk cost investments across a large base of economic activity.

ISVs specialize in developing applications that extend the functionality of the platform and add value to platform adopters, often in areas where the platform owner lacks expertise or where market conditions do not justify the platform owner's entry. ISVs have the option to become certified by SAP and become a member of the SAP platform ecosystem. This certification endorses the interoperability between the ISV's software and the SAP platform. In conjunction with SAP, the ISV undertakes development, documentation, and testing to ensure the product is compliant with SAP's platform specifications. Once the product successfully completes a certification test, a certification logo is issued by SAP, and the solution will be listed on SAP's Web portal which is accessible by its customers. The primary benefit to such partnering is to signal software compatibility and to give ISVs access and exposure to SAP's installed base. This may result in tangible financial benefits such as higher sales for ISVs (Ceccagnoli et al. 2012).

Although certification provides clear benefits for the ISVs, it also has the potential to increase the risk of entry by SAP. SAP (2005a) has made this potential for entry clear to ISVs, stating on its ecosystem Web portal that "Part of being an open ecosystem is open and fair competition among partners, and between SAP and partners. SAP cannot guarantee exclusivity of individual partner solutions, nor can we guarantee that we won't offer competing solutions." Competition from SAP could take several forms. For example, SAP could enter the complementary market and offer a directly competing product, or it could absorb some features of complements into one of its existing modules and therefore make it part of the platform. These different forms of competition are inherently difficult to observe and motivate our empirical approach: rather than examining the implications of appropriability mechanisms for SAP behavior, we instead examine their implications for ISV partnership decisions.

This risk of entry and the use of IPR to defend against it can be motivated by casual empirical evidence. One example is AMC Technology, a leading provider of multichannel integration solutions that allows contact centers to more efficiently manage all types of customer interactions. AMC Technology has been a certified SAP software partner since 1998. With its introduction of the product suite mySAP CRM 5.0 in 2005, SAP entered into AMC's market with a CRM Interaction Center module. SAP's new module allegedly contained copyrighted AMC code from AMC's Multi-Channel Management Suite (MCMS) product. AMC soon filed a lawsuit that claimed vicarious copyright infringement, breach of contract, and appropriation of trade secrets by SAP (SAP 2005b, Shapiro 2005). The U.S. district court awarded a preliminary injunction preventing SAP from "describing or purporting to authorize the copying, migration, or incorporation of AMC MCMS code" (Shapiro 2005, p. 27).

4. Methods and Measures

4.1. Sample

In this study we use the CorpTech directory of technology companies as the starting point to define our

[5] The distinction between appropriability mechanisms based on IP protection versus ownership of downstream capabilities is critical in this case, because growth is not necessarily associated with an increase in the marginal benefit of downstream capabilities. This is because whereas growth increases the threat of entry by the platform owner and thus the value of downstream capabilities, it is also associated with the earlier phases of the product life cycle, where downstream capabilities are a less important competitive factor. Put differently, downstream capabilities are more valuable in more mature phases of a product life cycle where sales growth is relatively *slower*, standards are established, and competition is focused on service, upgrades, and specialized marketing and manufacturing assets (Teece 1986).

sample. This data set has been used by scholars to study the value of patents in the software industry (Cockburn and MacGarvie 2009, Hall and MacGarvie 2010). In particular, we select the universe of firms—as defined by CorpTech—that produce software in categories that are likely to be complementary to an enterprise software platform and then to use variance in firm and product market characteristics among firms within those categories to identify the trade-offs shaping the partnership decision.

A critical challenge for our analysis is to identify the set of potential SAP partners. To do this, we utilize the CorpTech classification, which allows us to identify software firms and the types of products they offer. These "SOF" (software) codes have been used by prior researchers to identify market entry in software (Cockburn and MacGarvie 2009, 2011). We use these SOF codes to identify the set of firms that produce enterprise software and therefore face the decision to partner with SAP.

To identify the set of firms at-risk of partnering we choose the SOF codes with the highest propensity to partner with SAP. Using the complete list of SAP partners obtained from the SAP website, we find 411 U.S.-based software firms that are existing SAP partners. Comparing this list to the CorpTech directory generates 206 matching records. We then retrieve the distinct two-digit SOF codes of the 206 matching partners and identify the most frequent codes in their product portfolios. Because two product codes, SOF-MA (Manufacturing) and SOF-WD (warehousing and distribution), emerge as the most frequent, we use these as the starting point to identify our potential universe of ISVs.[6] We define our initial sample as the set of all CorpTech firms that have ever produced software products in the SOF-MA and SOF-WD categories between 1996 and 2004.

Although we focus on firms with products in two two-digit SOF codes, we note that the product portfolios of these firms extend far beyond these two categories. For example, among our final sample of 1220 ISVs, 474 also produce accounting software, 323 provide utility systems software, and 256 also provide sales/marketing software. The average number of SOF product codes that firms in our sample produce in is 3.54. Firms in these two categories represent 51% of the total number of SAP partners that we could identify in CorpTech. Because we are primarily interested in the commercialization strategies of start-ups, we only include ISVs established after 1980, with sales less than \$500 million, and with fewer than 1,000 employees throughout our study period.[7]

We restrict our sample period to 1996–2004. We begin our sample in 1996 because we find no partnership activities between SAP and start-up ISVs prior to 1996. The year 2004 represents the last year in our CorpTech database. Our final sample consists of 1,220 ISVs with 6,498 observations. The numbers of unique ISVs in the sample varies from 595 in the first sample year to 728 in the last.

4.2. Variable Definition and Operationalization

4.2.1. Dependent Variable. The dependent variable is whether an ISV enters into partnership with SAP in a particular year. We identify partnership formation events through press releases by searching LexisNexis.[8] For ISVs with multiple SAP partnership events (because of certification for multiple products, new product versions, or different interface certifications), we use the first such event as the time the ISV joins SAP's platform.

The unit of observation in our data is a firm-year, with the partnership variable equal to 1 if a first-time alliance is formed in that year and 0 otherwise. We do not expect appropriability mechanisms to have a significant impact on the length of partnership; in fact, all partners in our data remained partners throughout our sample period after they joined. Thus, the focus of our analysis is on the initial partnership decision, rather than the length of partnership, and therefore we delete postpartnership observations because the firms are no longer exposed to the hazard of forming a partnership with SAP.[9] In total, 35 of the 1,220 ISVs joined the SAP platform over the 1996–2004 period.

4.2.2. Independent Variables. *Patents and copyrights.* One of the major goals of this paper is to study how the possession of formal appropriability

[6] To verify that the unmatched partners are not systematically different from those matched to CorpTech, we collected information on the unmatched ISVs from Company Insight Center, a database owned by *BusinessWeek* and Capital IQ. From this database we obtained a short business profile for each of the remaining ISVs, complemented by a description of businesses and products collected from the ISVs' websites. Close examination of the profiles and product descriptions suggests that manufacturing software and warehouse/distribution software are also the two most frequently produced product categories by these unmatched ISVs, similar to the ISVs that are matched in the CorpTech database.

[7] These thresholds have also been used in prior studies of small firms (Petersen and Rajan 1994, Puranam et al. 2006). We also explored the use of alternative size thresholds and our results are robust to these changes. As an additional check, we visited each company website to confirm that the ISVs indeed produced enterprise software applications, and deleted those that did not. If the company no longer exists, we visited the archival website http://www.archive.org instead.

[8] To test the viability of this approach, we compared the list of partners obtained through this method with a list obtained from the SAP website and found that our method identifies 98% of partners mentioned by SAP.

[9] That is, we treat partnership as an absorbing state. We collected information on the status of ISVs after partnership and verified that this is the case.

mechanisms like patents or copyrights affect the decision to partner. Prior work examining the effect of court decisions that weakened copyright protection for software has shown that software firms use patents and copyrights as substitute appropriability mechanisms (Lerner and Zhu 2007). As a result, including these variables separately may miss important interactions between them. Thus, we compute a combined measure of patent and copyright use for our measure of the ISV's use of formal appropriability mechanisms.

To compute this combined variable, we first generate a measure of the stock of patents by date of grant using the U.S. Patent and Trademark Office (USPTO) patents database. Some vendors in our sample may have inventions in related areas such as IT hardware that will do little to deter SAP entry. Accordingly, we restrict our patent measure to software patents only. Identifying software patents is inherently difficult: because software is embedded in many products, there is no set of USPTO classes that maps to software inventions in the way that there is for other types of inventions. Nonetheless, one approach to identifying software patents is to use USPTO class-subclass combinations (Graham and Mowery 2006, Hall and MacGarvie 2010), whereas another is to use a Boolean query that searches for keywords that identify software inventions in the patent text (Bessen and Hunt 2007). These approaches have different advantages in identifying software patents and mitigating false positives (Hall and MacGarvie 2010). We take the intersection of these two approaches, as in Hall and MacGarvie (2010), to mitigate the effect of false positives on our data; however, our results are robust to alternative approaches.

Cockburn and MacGarvie (2011) demonstrated that citation-weighted patents have an economically strong and statistically significant impact on entry deterrence above and beyond the impact of patents per se; they argue that such "larger" patents may both be more difficult to invent around and represent more significant innovation by the inventor—both of which deter entry. Thus, to account for the heterogeneity in the size of prior patents and control for heterogeneity in the importance of the ISV's innovation, we use the weighted stock of patent grants by incorporating forward patent citations, using the well-known method of citation-weighting proposed by Hall et al. (2002). We obtain the firm's stock of registered copyrights from the U.S. Copyright Office. We retrieve the complete set of copyrights that are described as "computer files" within that office's classification scheme.

We define an aggregate measure of formal IPR protection, *High IP*. We set *High IP* equal to 1 if an ISV has either a high number of patents (greater than 0, because 95% of ISVs have zero patents) or a high number of copyrights (greater than the median). This combined measure is motivated in part by prior work on the commercialization strategies of start-up firms, which has employed dummy variables to measure multidimensional appropriation strategies that combine the use of patents and copyrights (Gans et al. 2002). As a robustness check, we also estimate regressions with separate continuous measures of patents and copyrights. Our results are robust to these alternative measurement strategies.

Downstream capabilities. Trademarks facilitate consumer choice among experience goods and transmit quality signals for infrequently consumed goods (Economides 1988). Although trademarks per se may not directly protect a firm against imitation, they enhance a firm's appropriability of its inventions by legally protecting its investments in marketing and other intangibles such as brand and reputation (Fosfuri et al. 2008). For example, "The Best-Run Businesses Run SAP" (U.S. trademark 78487112, owned by SAP AG) and "Global Access to Local Knowledge" (U.S. trademark 78655545, owned by Microsoft Corporation). We follow prior research on software-producing markets that has used trademarks as a proxy for the stock of marketing-specific downstream capabilities and a firm's brand capital (Fosfuri et al. 2008, Gambardella and Giarratana 2006). Brand capital is not easily contracted for through the market on competitive terms and represents a specialized complementary asset because it is hard to redeploy to alternative uses and by alternative users (Williamson 1991). We obtained the data from the USPTO trademarks database. We use only software trademarks that are currently "live" as of the date of observation. As we did for patents and copyrights, we define a discrete measure of this variable, *High trademark*, which is set to 1 if an ISV's stock of trademarks is greater than the sample median and 0 otherwise. An alternative approach would be to simply use the count of trademarks. The advantage of this latter approach is that it would more precisely capture the effects of the intensive margin of trademarks on the partnership decision; the disadvantage is that the distribution of trademarks is highly skewed, and using the count requires us to make an assumption on the likely nonlinearity of their effects. We have experimented with alternative models that use the log of the number of trademarks (together with patents and copyrights) and our qualitative results remain supportive of Hypotheses 1 and 2.

Alternative measure of downstream capabilities. Although a firm's stock of trademarks may be an adequate proxy for its marketing capabilities, we acknowledge that the downstream capabilities of an ISV may encompass other relevant dimensions not captured by the firm's stock of trademarks, such as its consulting and other professional service capabilities. As a robustness check of our measure of marketing capabilities, we construct a broader measure that combines both marketing and software consulting services capabilities. The CorpTech database provides information on each firm's portfolio of software service offerings. These service offerings describe a firm's software consulting services, which are particularly relevant to the enterprise software industry. The discrete variable, *High downstream*, is set to 1 if an ISV owns a high number of trademarks or if it provides any software consulting services.

Sales growth. We construct two variables measuring the growth prospects of each ISV based on alternative ways of characterizing growth—demand in vertical industry and size of horizontal product market. Whereas the former highlights the risk of transferring an ISV's industry-specific know-how and demand signal, the latter emphasizes the risk of transferring knowledge related to product designs and functions. Our first variable, *Sales growth*, is computed based on the vertical industry segments that an ISV serves instead of the products that it sells because most firms sell only to one or two industries and sell a larger number of software products, and CorpTech records ISV sales at the firm level instead of at the product level. That is, we use industry demand data for software to identify growth opportunities of the ISVs that serve the industry. To do this, we use the target industry descriptions of the ISVs in our CorpTech data. These target industry descriptions describe the client industries that a firm sells its products and services to. We read the descriptions and manually code them using the SAP "master code" industry classification system, which is composed of indicators for 33 vertical industries (e.g., aerospace and defense, banking, and chemical, to name a few). The sales growth rate in each industry-year (or master code-year) is calculated using the CorpTech universe of ISVs, and is equal to (*current year sales of all ISVs that serve the industry*)/(*previous year sales of all ISVs that serve the industry*). The firm-year growth variable, *Sales growth*, is defined as the arithmetic average of growth rates across all the industries that an ISV serves.

Entry rate. We use the entry rate in an ISV's product market as a second measure of growth opportunities. This variable serves as a proxy for product market growth for two reasons. First, entry is likely to be correlated with sales growth because, ceteris paribus, larger markets will support more firms. Furthermore, entry is correlated with the stage of the product life cycle (e.g., Klepper 1996, Utterback and Abernathy 1975): industries with high entry rates tend to be younger and have significant growth opportunities. We treat each CorpTech SOF product code as a distinct market and count the number of firms that produce the product in each year. Market (or product code) level entry or exit rate is defined as (*number of ISVs in the current year* − *number of ISVs in the previous year*)/(*number of ISVs in the current year*).[10] The variable *entry rate* is subsequently defined as the arithmetic average of market level entry rates across the ISV's product portfolio.

4.2.3. Control Variables. *SAP installed base.* To construct a valid measure of SAP's market penetration in the ISV's target industries, we obtain SAP installation data in the United States from the Harte-Hanks CI Technology Database. Harte-Hanks surveys over 300,000 establishments in the United States per year on their use of information technology; our sample of data from the CI database includes all establishments with over 100 employees. We identify the set of firms that have adopted SAP in each year and weight these by the number of employees.

We use these data to compute the SAP penetration rate for the industries served by the ISVs. As we did for sales growth, the client industries of the ISVs are coded using the SAP "master code" industry classification system. Using the data on installed base, we calculate the employee-weighted penetration rate in the CI universe of firms for each industry-year.[11] The variable *target industry penetration* is defined as the average SAP penetration rate across all the industries that an ISV serves. In short, this variable measures the penetration of SAP within the downstream industries to which the ISV sells its product.

Product overlap with the platform owner. We use the similarity in product market space between the ISV and SAP to control for the potential effects of product competition between the two. We retrieve the distinct SOF product codes from CorpTech for each ISV in each year, and compare those with SAP's SOF product code portfolio in the same year from CorpTech. SOF product codes are used as a proxy for product lines because they correspond very well to the functional modules of enterprise software. The variable *product_overlap* is defined as the ratio of the number of common product codes (produced by both an ISV and SAP) to the total number of an ISV's product codes for each firm-year. There are 2,168 of 6,498 firm-year observations in our sample for which there is zero product overlap. This measure reflects a number

[10] We use number of ISVs in the current year as the denominator in our definition of entry rate to avoid divided-by-zero issues that arise when the number of ISVs in the previous year is zero, leaving this variable undefined. Our results are robust to using number of ISVs in the previous year as the denominator and dropping observations where this variable is undefined.

[11] As the firms in CI database are identified with Standard Industrial Classification (SIC) codes instead of SAP's master codes, this step involves creating a mapping table between SIC codes and master codes.

of factors, including changes in SAP's product space over time and an ISV's product market entry and exit decisions.

Publications. As a control for an ISV's innovative culture and research output, we obtain the ISV's cumulative number of *publications* in academic journals or conferences in each year via the Web of Science database. To account for the importance of publications, we also retrieve forward citation data for all the publications and construct citation-weighted publications.

Product innovations. To check the robustness of our analysis to the inclusion of variables capturing the level and quality of an ISV's software product offerings, we also include a control for the number of new products and product versions offered by the ISV. A similar measure was used recently to proxy for a firm's innovative performance by Fosfuri et al. (2008). To construct this variable, we collected all press releases and articles on product introductions for all the ISVs in our sample. These data were obtained from *Business and Company Resource Center*, a database from Gale. We gathered all the articles under the class "products and services," read the articles, and coded them as either a new product introduction or a new version update. The variable *product innovation* is defined as the cumulative sum of all new product and version introductions. We also experimented with including a variant of this variable that includes only new product introductions (and excludes new versions), and the results were qualitatively similar.

County employment. To control for the effect of local market characteristics on partnership formation, the location (zip code) of each ISV's headquarters was identified from the CorpTech database, which was in turn used to identify the county where the ISV is located. We then obtained county-level employment data from U.S. Census County Business Patterns data and derived the variable *county employment* as the sum of local employment in an ISV's county of residence.

Other controls. We control for various firm-level drivers that could influence an ISV's decision to join the platform. Firm size is measured by an ISV's number of *employees*, obtained directly from the CorpTech database.[12] ISV *age* is derived by referencing the year that an ISV was established, according to its record in the CorpTech database. To allow for nonlinear effect of age, we add both linear and quadratic terms. We also include controls for firm funding sources, because an ISV's source of capital is likely to affect its decision to form partnerships (Colombo et al. 2006, Gans et al. 2002). We create three dummy variables, *corporate investment*, *private investment*, and *venture capital* (*VC*) *investment*, corresponding to the funding sources of the ISVs as categorized by the CorpTech Database.

Table 1 provides summary statistics of the variables and controls, as well as the comparison between nonpartners and those firms that eventually became partners with SAP during our sample period. It is worth noting that *patents* are far less frequently used by start-up ISVs in the enterprise software industry (with a mean of 0.07 patent per firm) than *copyrights* (a mean of 1.98 per firm), consistent with prior literature suggesting that copyrights remain an important source of IP protection for enterprise software because most innovations are in business processes, routines, and best practices that may not be patentable (Mann and Sager 2007).

4.3. Model Specification

In our baseline analysis we use hazard models to analyze how the presence of formal IPR mechanisms and downstream capabilities shape the time to partnership with SAP. The hazard model (also referred to as survival, duration, or event history model) is a useful approach for our setting because it directly models time to event, relaxes the normality assumption imposed in linear regression (the data generation process of time-to-event data usually produces a skewed distribution in the error component rather than a symmetric one), and provides an approach to address the incomplete observation of survival times when censoring occurs (Hosmer and Lemeshow 1999). Hazard analysis models the underlying and unobserved hazard rate, which is the instantaneous rate at which hazard events occur at time t, given that the subject under study has survived until time t.

We chose the Cox proportional hazard model as our baseline specification. This model is a semiparametric specification that makes no assumptions about the shape of the baseline hazard over time and assumes that covariates multiplicatively shift the baseline hazard function. In our benchmark specification, we estimate $h_i(t \mid \mathbf{x}_{i,t-1}) = h_0(t)\exp(\mathbf{x}_{i,t-1}\boldsymbol{\beta})$, the conditional instantaneous hazard rate for ISV i in year t, with $h_0(t)$ being the unspecified baseline hazard in year t, and

$$\begin{aligned}\mathbf{x}_{i,t-1}\boldsymbol{\beta} = {} & \beta_0 \mathit{High\ IP}_{i,t-1} + \beta_1 \mathit{High\ trademark}_{i,t-1} \\ & + \beta_2 \mathit{High\ IP}_{i,t-1} \times \mathit{High\ trademark}_{i,t-1} \\ & + \beta_3 \mathit{Sales\ growth}_{i,t-1} + \gamma \mathbf{Z}_{i,t-1}; \end{aligned} \tag{1}$$

$\mathbf{Z}_{i,t-1}$ represents a vector of time-varying firm, industry, and location control variables, all lagged by one year to allow for their delayed effects on partnership

[12] We have experimented with including a control for firm sales, however we exclude this variable from our baseline results because the high correlation (> 0.9) between sales and employees. Our results are robust to the inclusion of the sales variable, however.

Table 1 Summary Statistics

	Whole sample				Nonpartner				Partner				Differences
Variables	Mean	Std. dev.	Min	Max	Mean	Std. dev.	Min	Max	Mean	Std. dev.	Min	Max	Nonpartner mean minus partner mean
Patents	0.067	0.377	0	6	0.065	0.369	0	6	0.192	0.665	0	3	-0.127^{***}
Copyrights	1.978	13.022	0	498	1.964	13.065	0	498	2.725	10.489	0	80	-0.761
Trademarks	0.808	2.004	0	23	0.774	1.946	0	23	2.650	3.611	0	15	-1.876^{***}
SAP penetration	0.228	0.154	0.001	0.932	0.228	0.155	0.001	0.932	0.260	0.117	0.080	0.582	-0.032^{**}
Entry rate	0.042	0.069	−0.231	1	0.041	0.068	−0.231	1	0.079	0.097	−0.086	0.558	-0.038^{***}
Sales growth	1.261	0.338	0.873	5.637	1.261	0.339	0.873	5.637	1.268	0.300	1.000	2.939	-0.007
SAP product overlap	0.422	0.326	0	1	0.424	0.326	0	1	0.269	0.298	0	1	0.155^{***}
Public	0.054	0.226	0	1	0.051	0.220	0	1	0.200	0.402	0	1	-0.149^{***}
Age	12.935	5.693	0	24	12.993	5.680	0	24	9.858	5.560	1	22	3.134^{***}
Employees	51.794	93.743	1	900	49.451	89.719	1	900	176.350	180.237	7	900	-126.899^{***}
Corporate investment	0.042	0.200	0	1	0.041	0.197	0	1	0.092	0.290	0	1	-0.051^{***}
Private investment	0.520	0.500	0	1	0.525	0.499	0	1	0.283	0.453	0	1	0.242^{***}
VC investment	0.114	0.318	0	1	0.108	0.311	0	1	0.425	0.496	0	1	-0.317^{***}
Publications	0.608	5.484	0	137	0.610	5.533	0	137	0.475	1.174	0	6	0.135
Product innovations	0.248	0.963	0	15	0.231	0.936	0	15	1.117	1.696	0	8	-0.885^{***}
County employment	586,810	626,472	1,490	3,548,191	583,805.9	625,218.3	1,490	3,548,191	746,480	673,548.2	14,656	3,548,191	$-162{,}674^{***}$

Note. N (full sample) = 6,498; N (nonpartner) = 6,378; N (partner) = 120.
$^{**}p < 0.05$; $^{***}p < 0.01$.

decisions. To test Hypothesis 3 we include in the conditional instantaneous hazard rate the interactions of *Sales_growth* with the appropriability variables.

Because we use nonlinear models with interactions, we cannot test our hypotheses by directly examining the sign of our coefficients (Ai and Norton 2003). In particular, in our baseline model, we study the effect of a change in a variable $x_{i,t-1}$ on $\log h(\mathbf{x}_{i,t-1}\boldsymbol{\beta})$, i.e., we examine the semielasticities of the hazard rate with respect to a change in each of the key independent variables of interest. Thus, because the variables are discrete, Hypothesis 1 is tested based on estimates of the following semielasticities:

$$\log h(\mathbf{x}_{i,t-1}\boldsymbol{\beta})\,|_{High\,IP_{i,t-1}=1} - \log h(\mathbf{x}_{i,t-1}\boldsymbol{\beta})\,|_{High\,IP_{i,t-1}=0} > 0 \quad \text{and}$$

$$\log h(\mathbf{x}_{i,t-1}\boldsymbol{\beta})\,|_{High\,trademark_{i,t-1}=1} - \log h(\mathbf{x}_{i,t-1}\boldsymbol{\beta})\,|_{High\,trademark_{i,t-1}=0} > 0.$$

Hypotheses 2 and 3 are tested similarly.

4.4. Addressing Alternative Explanations

Table 1 demonstrates that there are significant differences in observable characteristics of firms that eventually partner and those that do not in the pooled sample: in addition to differing appropriability mechanisms, partners have higher SAP penetration rates and lower product overlap with SAP, are younger, and are larger. One potential concern is that there exist unobserved differences among firms that are correlated with appropriability mechanisms and partnership decisions. If such unobservables exist, then we will be unable to identify the causal relationship of interest. One particular concern is that appropriability mechanisms like patents and copyrights may be correlated with unobserved firm quality, or with differences in the type of product developed. We address this concern—and others related to unobserved quality—through several means, as outlined below.

4.4.1. Direct Measures of Firm Product Quality. In addition to variables controlling for firm quality in the baseline, such as a firm's publications, we also add a control for new or improved product introductions. This variable will help us to control for cross-sectional and time-varying differences in the new product intensity of firms in our sample that may be correlated with the likelihood of partnership.

4.4.2. Binary Response Panel Data Models with Unobserved Firm Heterogeneity. To control for time-invariant unobserved firm heterogeneity, we run additional robustness checks using firm fixed effects (FE) models. Of course, estimating nonlinear models such as hazard models using firm fixed effects is likely to lead to biased and inconsistent estimates because of the well-known incidental parameters problem. We address this problem through two approaches that model directly the discrete choice of whether to partner, rather than modeling the hazard rate. In particular, with reference to testing Hypotheses 1 and 2, we directly estimate

$$\begin{aligned}\mathrm{Prob}(Partner_{i,t}) &= F(\beta_0 High\ IP_{i,t-1} + \beta_1 High\ trademark_{i,t-1} \\ &\quad + \beta_2 High\ IP_{i,t-1} \times High\ trademark_{i,t-1} \\ &\quad + \beta_3 Sales\ growth_{i,t-1} + \gamma \mathbf{Z}_{i,t-1} + \mu_i), \qquad (2)\end{aligned}$$

where $Partner_{i,t}$ is a binary variable indicating whether firm i partners with SAP in time t; $\mathbf{Z}_{i,t-1}$ represents a vector of time-varying firm, industry, and location control variables; and μ_i is a time-invariant unobserved effect.

In our first model, we assume $F()$ is the cumulative distribution function of a normal distribution. An increasingly popular method to estimate binary response panel data models that accounts for within-firm nonindependence of observations and unobserved firm-specific effects is the generalized estimating equation (GEE) approach (Wooldridge 2001, Zeger et al. 1988). We employ the device used by Mundlak (1978) and Chamberlain (1980) of directly modeling the conditional distribution of the unobserved effects, assuming the mean of this distribution is a linear combination of the means of the complete set of our independent variables. The regression coefficients from this model provide information about the average response across firms ("population averaged") rather than how one firm's response changes with the covariates (Zeger et al. 1988). With reference to our study, for example, this model provides an estimate of the differential rate of partnering with the platform for firms with or without formal IPR protection.

Our second approach to adding firm fixed effects is a linear probability model (LPM). The LPM can be viewed as a linear approximation to the nonlinear model expressed above in Equation (2). It is well known that the LPM has its limitations—in particular, it can predict probabilities outside a zero–one interval, and its error term is inherently heteroskedastic. However, with the appropriate robust standard error corrections, this model can provide useful approximations to the underlying relationship of interest, so long as they are not used to predict too far out of sample (Angrist and Pischke 2009). Furthermore, prior work has shown that the LPM generates reasonable estimates within the region of support of the data (e.g., Miller and Tucker 2009).

4.4.3. Fixed Effects Linear Probability Model with Instrumental Variables (IV). Although we address time-invariant unobserved heterogeneity in our binary response models, another concern is the potential correlation between a firm's appropriation strategies and unobservable changes over time to firm, location, or industry characteristics. To address these concerns, we examine the robustness of our LPM results using changes in software patent strength to instrument for *High IP*.

During and immediately preceding our sample period, there were several changes to legal regimes that clarified and strengthened the patentability of software inventions (Cockburn and MacGarvie 2009, 2011; Hall and MacGarvie 2010). Hall and MacGarvie (2010) provided a detailed accounting of these changes. Prior to 1996, patent protection was understood to be limited to software directly tied to physical processes such as manufacturing.[13] However, a series of court cases in 1994 and 1995 resulted in the USPTO issuing in 1996 a set of new guidelines on the patentability of software. Other authors have noted that these changes in regime were associated with a significant increase in the volume and growth of patenting in related software categories, with "treated" categories seeing faster rates of growth after the regime change (Hall 2009, Hall and MacGarvie 2010). Furthermore, indirect evidence of the implications of the regime change can be seen in the growth in institutions to mitigate the patent thicket problem in software, such as the growth of intellectual property disclosures through standard-setting organizations (Rysman and Simcoe 2008) and the growth of patent pools in related areas (e.g., Hall and Helmers 2011, Wen et al. 2011).

This variation is the source of our instrumental variables strategy. As has been noted elsewhere in this paper, there remained considerable uncertainty about the patentability of software during and immediately preceding our sample period. Changes in legal regimes that increase the likelihood that a software patent will be upheld in court will increase the benefits of software patenting and should increase the propensity of firms in our sample to patent software. Other things equal, this should result in an increased likelihood of the use of formal IPR, and an increased likelihood of observing $High\ IP = 1$.[14]

To construct our instruments, we assign CorpTech SOF product codes to one of two classes based upon the patents in those product markets. These classes reflect the characteristics of patents and when they were affected by the regime changes described above: The first set of product codes includes both those with patents that had strong appropriability prior to the regime change (these are largely patents that operate on physical media; Cockburn and MacGarvie 2011) and those with patents that had weak appropriability before and after the regime change (e.g., business method patents). The second set of "treated" product codes are those with patents that were strengthened after the regime change described above. To identify which product codes were treated, we use the patent-to-product code classification described by Cockburn

[13] This was specified in the 1981 Supreme Court decision *Diamond v. Diehr*.

[14] If patents and copyrights are substitutes, then an increase in appropriability strength of one may decrease the use of the other. So long as this substitution is not one to one, increases in patent strength should still lead to an increase in the likelihood of observing $High\ IP = 1$. We experimented with regressions of the log of ($1 +$ number of copyrights) on changes in the legal regime and a set of controls, and found that changes in regime status had no statistically significant impact on the number of copyrights.

and MacGarvie (2009, 2011).[15] For each ISV, we created a dummy indicating whether the firm operated in a product code that was treated by the regime change. We then interacted this treatment group dummy with a time dummy used to indicate the timing of the change in patent strength. Whereas the timing for the regime-shifting events in our data occurred in 1996, other authors have found that the timing of delay for software patents between the patent application and grant was 2.8 years over much of our sample period (Cockburn and MacGarvie 2011), meaning that the earliest these regime changes would influence patent propensity would be 1999. We further interact each of these instruments with a *High trademark* dummy to address the potential endogeneity of *High IP* × *High trademark*. In all, we have two potentially endogenous variables in our regressions, *High IP* and *High IP* × *High trademark*, and four instruments, a dummy for the treatment group, the interaction of the treatment group with a 1999 dummy, the interaction of the treatment group with *High trademark*, and the three-way interaction of treatment group, *High trademark*, and the 1999 dummy.

The key assumption for our IV strategy is that changes in patenting regime will be uncorrelated with changes in firm-level unobservables such as firm quality, resources, or product choice. To the extent this assumption is plausible, our instrumental variable strategy—in conjunction with our other controls and robustness checks—increases confidence in our interpretation that increases in formal IPR will lead to an increased likelihood of partnership through stronger appropriability. Unfortunately, we were unable to identify a similar instrument that would shift the likelihood of observing *High trademark*. However, to the extent that we are able to control for a variety of observable measures of firm-level quality, our estimates for *High trademark* are still informative about the types of firms most likely to engage in partnership.

4.4.4. Additional Survival Analysis. We conduct further robustness checks to probe the distributional assumptions of our Cox model. Although the Cox hazard model assumes a continuous-time hazard rate function, often the survival times are not observed more precisely than the interval within which the event occurred. In our setting, partnership events are observed within the interval of a year. Discrete-time hazard models are often employed to investigate the relationship between interval-censored survival time and a set of explanatory variables. Particularly, Prentice and Gloeckler (1978) showed that if the data are generated by a continuous-time proportional hazard model, coefficients estimated by a binary response model with a complementary log–log link function are equivalent to those of the continuous-time proportional hazard model. As an alternative strategy, we also present results from a complementary log–log estimation of the event histories of SAP partnerships.

Finally, although hazard models are typically used to analyze "time-to-event" data, they often assume that the event in question is inevitable in the sense that the probability of eventual failure is greater than zero for all individuals. In contrast, split population survival models (Schmidt and Witte 1989), labeled *cure models* by biostatisticians, suppose that a proportion of the sample never fails, and therefore is immune to the event.[16] These models explicitly estimate the fraction of the immune population, as well as the parameters characterizing the hazard rate for the rest of the population (usually by a discrete-time proportional hazard model such as complementary log–log). We present the results of a split population survival model using the complementary log–log model as a robustness check.

5. Results

Before we present the results from the empirical models, in Table 2 we provide some motivational statistics. Table 2(a) shows the conditional probability of partnering for each firm at the time of entry and exit from our sample, and depending on whether it goes from low to high IP, remains at low IP, or remains at high IP throughout the sample. These statistics highlight the variance in our data that identifies the association between IPR and the likelihood of partnership. Among the 76 firms that go from low to high IP, the likelihood of partnership increases by 7.9 percentage points during our sample; this compares to a 1.6 percentage point increase for those firms for which IP stays low (867 firms), and a 2.5 percentage point increase for those for which IP stays high (277 firms). A t-test indicates that the growth in likelihood of partnership during our sample is significantly greater for firms switching from low to high IP than for firms remaining at low IP (significant at the

[15] In our data, the following SOF product codes were in Classes 1 and 2. Class 1 (unaffected by the regime change) included manufacturing, accounting, banking, construction, educational/training, financial analysis/management, government, health services, insurance, legal, library, nonprofit organization, natural resource management, project management, public utilities, real estate, sales/marketing, service industry, transportation, warehousing/distribution, and applications software not elsewhere classified. Class 2 (affected by the regime change) included artificial intelligence, communications system, database/file management, media communications, office automation, program development, technical/scientific, and utility systems.

[16] We thank an anonymous reviewer for suggesting split population survival analysis.

Table 2 Likelihood of Observing a Partnership at the Beginning and End of the Sample, By Firm

(a)				(b)			
IPR	(1) First observation in sample	(2) Last observation in sample	(3) Difference between first and last observation	Trademarks	(4) First observation in sample	(5) Last observation in sample	(6) Difference between first and last observation
Changes from low to high IP	0.0000 ($N = 76$)	0.0789 ($N = 76$)	0.0789** ($N = 76$)	Changes from low to high trademark	0.0000 ($N = 175$)	0.0457 ($N = 175$)	0.0457*** ($N = 175$)
IP stays low	0.0046 ($N = 867$)	0.0208 ($N = 867$)	0.0161*** ($N = 867$)	Trademark stays low	0.0064 ($N = 782$)	0.0153 ($N = 782$)	0.0090*** ($N = 782$)
IP stays high	0.0144 ($N = 277$)	0.0397 ($N = 277$)	0.0253*** ($N = 277$)	Trademark stays high	0.0129 ($N = 232$)	0.0560 ($N = 232$)	0.0430*** ($N = 232$)
Changes from high to low IP	N/A ($N = 0$)	N/A ($N = 0$)	N/A ($N = 0$)	Changes from high to low trademark	0.0000 ($N = 31$)	0.0000 ($N = 31$)	0.0000 ($N = 31$)

Note. Cells represent the likelihood of observing a partnership dependent on the firm's entry/exit year and changes in the firm's IP and trademark status throughout the sample.
$^{**}p < 0.05$; $^{***}p < 0.01$.

1% level) or those remaining at high IP (significant at the 5% level). Similarly, examining the results in column (2), the mean probability of partnership is higher for those with high IP compared to those with low IP. Overall, the results in Table 2(a) suggest that both between-firm and within-firm variation in IP in our data will identify the association between IP holdings and the likelihood of partnership.

Table 2(b) presents a similar set of statistics for trademarks. The likelihood of partnership increases by 4.6 percentage points over our sample for those firms that go from a low number of trademarks to a high number of trademarks (175 firms), increases by 4.3 percentage points for those that have a high number of trademarks throughout our sample (232 firms), and increases by 0.9 percentage points for those for which the number of trademarks stays low throughout our sample (782 firms). The change in the likelihood of partnership is significantly greater for those firms that switch from a low to a high number of trademarks compared to those that remain at a low number of trademarks (at the 1% level); however, there is no significant difference between those that switch to a high number of trademarks and those that remain at a high number of trademarks throughout the sample. Examination of column (5) shows that the mean likelihood of partnership is significantly higher for those firms with a high number of trademarks compared to those without. Overall, the results in Table 2(b) suggest that between firm variance will primarily identify the effect of trademarks in our data. We speculate that this result could reflect increases in the value of trademarks as an appropriability mechanism over time; however, we admittedly have no way of testing this conjecture.

5.1. Tests of Hypotheses 1 and 2

Results from the Cox proportional hazard model, complementary log–log model, and split population survival model that investigate the effects of IPR, trademarks, and their interactions on partnership decisions are presented in Table 3. For each model, we calculate the semielasticities (ey/dx) of major independent variables. Column (1) presents the baseline model where discrete measures of IPR, trademarks, and their interactions are included. In column (2) we add the full set of control variables. In column (3) we replace the *High trademark* variable with a broader measure of downstream capability, *High downstream*, which also encompasses the consulting service capabilities of the ISVs. In column (4) we add *Product innovation* as a control for heterogeneity in the innovativeness and quality of ISVs. Results from a complementary log–log discrete-time survival model are presented in column (5). Finally, in column (6) we show the estimates from the split population survival model that explicitly allows for a proportion of ISVs not to be at risk for experiencing the partnering event.

The results of the hazard models suggest that increases in IP and trademarks are associated with a higher likelihood of partnership across all specifications. We use the marginal effects at the bottom of Table 3 to measure the impact of these two variables. Possession of stronger IPR or stronger downstream capabilities by an ISV is associated with a greater likelihood of joining the SAP platform. This result is robust to the use of an alternative and broader measure of downstream capabilities that includes the ISV's consulting activities (column (3)), and holds when controlling for the ISV's new product innovations (column (4)).

Table 3 Hazard Models

Variables	(1) Baseline Cox	(2) Full set of controls	(3) Alternative downstream	(4) Includes product introductions	(5) Complementary log–log	(6) Split population cure model
High IP	1.300**	1.395**	1.371**	1.421**	1.398**	1.406**
	(0.539)	(0.553)	(0.554)	(0.556)	(0.596)	(0.549)
High trademark[a]	1.741***	1.092**	0.820*	1.080**	0.865*	1.009**
	(0.480)	(0.495)	(0.496)	(0.500)	(0.519)	(0.498)
High IP × *High trademark*[a]	−1.075	−1.356*	−1.229*	−1.420*	−1.293*	−1.412**
	(0.684)	(0.723)	(0.733)	(0.732)	(0.783)	(0.710)
Sales growth		−0.760	−0.787	−0.723	0.179	−0.600
		(1.295)	(1.285)	(1.265)	(0.700)	(0.781)
Entry rate		1.330	1.269	1.094	3.787***	2.288
		(1.435)	(1.457)	(1.525)	(1.052)	(1.587)
SAP penetration		2.141**	2.123**	2.007**	1.908**	1.716
		(0.871)	(0.840)	(0.881)	(0.879)	(1.107)
SAP product overlap		−0.838	−0.856	−0.794	−0.998*	−0.825
		(0.540)	(0.547)	(0.536)	(0.536)	(0.559)
Age		−0.051	−0.0447	−0.072	−0.090	−0.068
		(0.121)	(0.121)	(0.125)	(0.115)	(0.122)
Age^2		0.0009	0.0005	0.002	0.001	0.001
		(0.0054)	(0.005)	(0.006)	(0.005)	(0.005)
Log *Employee*		0.700***	0.711***	0.669***	0.716***	0.695***
		(0.142)	(0.138)	(0.141)	(0.152)	(0.160)
Corporate investment		0.256	0.251	0.296	0.127	0.199
		(0.623)	(0.636)	(0.624)	(0.686)	(0.622)
Private investment		−0.523	−0.576	−0.484	−0.594	−0.561
		(0.422)	(0.426)	(0.423)	(0.440)	(0.381)
VC investment		0.871**	0.900**	0.848**	0.848**	0.863**
		(0.391)	(0.385)	(0.400)	(0.411)	(0.391)
Log *Publication*		−0.007	0.002	−0.043	−0.006	0.001
		(0.244)	(0.240)	(0.249)	(0.229)	(0.346)
County employment		0.166	0.156	0.197	0.191	0.155
		(0.174)	(0.172)	(0.178)	(0.186)	(0.183)
Product innovation				0.138*		
				(0.075)		
Marginal effects	*ey/dx*	*ey/dx*	*ey/dx*	*ey/dx*	*ey/dx*	*ey/dx*
High IP (average)	0.985**	0.998**	0.922**	1.006**	1.017**	0.993**
	(0.404)	(0.422)	(0.401)	(0.427)	(0.457)	(0.414)
High IP (*High trademark* = 0)[a]	1.300**	1.395**	1.372**	1.422**	1.394**	1.406**
	(0.539)	(0.553)	(0.554)	(0.556)	(0.594)	(0.549)
High IP (*High trademark* = 1)[a]	0.224	0.039	0.142	0.001	0.105	−0.006
	(0.433)	(0.490)	(0.499)	(0.504)	(0.541)	(0.462)
High trademark (average)[a]	1.432***	0.701*	0.466	0.671	0.492	0.602
	(0.370)	(0.410)	(0.401)	(0.418)	(0.421)	(0.393)
High trademark (*High IP* = 0)[a]	1.741***	1.092**	0.820*	1.080**	0.863*	1.009**
	(0.480)	(0.495)	(0.496)	(0.500)	(0.518)	(0.498)
High trademark (*High IP* = 1)[a]	0.666	−0.264	−0.409	−0.340	−0.427	−0.403
	(0.489)	(0.593)	(0.578)	(0.610)	(0.620)	(0.529)

Notes. All regressions are estimated over time to SAP partnership. The number of observations in all specifications is 6,498 (unbalanced panel of 1,220 firms observed over a nine-year period). Columns (1)–(4) show the results of Cox proportional hazard models, column (5) uses a complementary log–log regression, and column (6) shows the results of a split population cure model. All marginal effects are semielasticities (and so are denoted by the notation *ey/dx*) of the hazard rate: they represent $\log h(\boldsymbol{x}_{i,t-1}\boldsymbol{\beta})|_{x_{i,t-1}=1} - \log h(\boldsymbol{x}_{i,t-1}\boldsymbol{\beta})|_{x_{i,t-1}=0}$ for discrete variables. Heteroskedasticity robust standard errors (clustered over firms) are in parentheses, except in column (6), which uses classical (independent and identically distributed) standard errors.

[a]Column (3) uses a broader measure of downstream capabilities that combines both marketing and software consulting services capabilities.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Consistent with Hypothesis 2, we find evidence that IP protection and downstream capabilities act as substitutes for one another in their influence on partnership formation. The interaction effect is negative and significant at conventional levels across models (a formal test based upon marginal effects is discussed below).

To understand the magnitudes of the main effects and interaction effects of IPR protection and trademarks, we present semielasticities at the bottom of Table 3. With reference to the baseline model with the full set of controls (in column (2)), we find that, on average, a discrete change of the *High IP* variable from 0 to 1 is associated with a 99.8% increase in the hazard rate of an ISV joining the SAP platform. This effect is significantly moderated by the ISV's ownership of downstream capabilities because of the significant interaction effect between *High IP* and *High trademark*. For example, while the effect of *High IP* is magnified when an ISV's ownership of trademark is low (139.5% increase in the hazard of partnering), the magnitude is much smaller when the ISV's ownership of trademarks is high (3.9% increase in the hazard rate) and is not significantly different from 0. Similar patterns are observed for the marginal effect of the *High trademark* variable. A change of *High trademark* from 0 to 1 is associated with a 70.1% increase in the hazard rate when evaluated at average values for patents and copyright—this effect is much larger when the ISV is not protected by patents or copyrights (109.2% increase in the hazard rate). On the other hand, the marginal effect of *High trademark* is not statistically significant if an ISV has a high level of IP protection. The patterns are similar when discrete-time hazard models are employed.

5.1.1. Robustness Tests. Table 4 presents the results of the GEE probability model, the fixed effects LPMs, and the fixed effects instrumental variable LPMs. Results of these models are consistent with those presented above, though these models have a more intuitive interpretation based on standard marginal effects on probabilities, rather than the marginal effect on hazard rates as in the previous subsection. ISVs with strong IPR have an approximately two percentage point higher probability of partnering, on average, across specifications. Though the size of the effect is modest in absolute terms, because the proportion of partnering firms is small, the effect relative to the baseline is actually large. The marginal effects of IPR evaluated at average *High trademark* are all statistically significant at the 10% significance level, except for the GEE model, where it is significantly different than zero at the 16% level. Although the results in Table 4 show that the impact of IPR can be identified using only within-firm variance in our data, we are unable to identify any effect from increasing trademarks using only within-firm variance. This result was foreshadowed in column (6) of Table 2: although increases in the number of trademarks throughout the sample are associated with an increase in the likelihood of partnership, this increase is only slightly higher than that observed for firms that remained at a high number of trademarks through the sample. In short, our results for the effect of downstream capabilities are identified using between-firm variance. Although these results may reflect unobserved differences in the value of brand or service across firms, we believe this interpretation is consistent with the hypothesis set forth in §2.

The marginal effect of *High IP* is much higher for the IV estimates: column (5) of Table 4 shows that the marginal effect of strong IP protection at the average level of downstream capabilities is equal to 29 percentage points and significant at the 10% level. We believe this result reflects a local average treatment effect: although our instruments are uncorrelated with unobserved factors that influence partnership, the marginal effect of *High IP* is greatest among those firms that are most influenced by the policy change. However, a Hausman test is unable to reject the null hypothesis that the parameter estimates in columns (2) and (4) and columns (3) and (5) are the same. The first stage regressions (not shown) indicate that the joint F-test of the set of excluded instruments is significant at the 1% level for both endogenous variables (the *High IP* and *High IP* $\times$ *High trademark* interaction), with F-statistics equal to 4.03 and 4.26. The test of the overidentification restrictions cannot reject the null that the exclusion restrictions are valid, supporting our IV strategy. The IV results in column (5) also support the existence of a substitution effect between *High IP* and *High trademark*, because the marginal effect of strong IP protection at low levels of trademarks (*High trademark* $= 0$) jumps to 34 percentage points (significant at the 10% level), whereas it drops to 17 percentage points (and not significantly different than zero) at high levels of trademarks (*High trademark* $= 1$).[17]

[17] We conducted a series of additional robustness checks that we do not include here to conserve space. We tested the robustness of our fixed effects linear probability model to alternative lag structures. One potential concern of our sample construction is that we may be dropping firms that grow too large (and that are potentially very successful) from our sample. We have estimated a model that includes firms that started small but were excluded from our baseline sample because they grew too large (this lead to an increase of 521 firm-year observations in the sample). We also tested the robustness of our hazard models to the use of a multiplicative firm-specific error term (or *frailty*). Our results are robust to all of these analyses.

Table 4 Probability Models

Variables	(1) GEE (probit link)	(2) LPM with FE, without interaction	(3) LPM with FE, with interaction	(4) LPM with FE and IV, without interaction	(5) LPM with FE and IV, with interaction
High IP	1.328*** (0.443)	0.017* (0.010)	0.028** (0.014)	0.288* (0.155)	0.338* (0.203)
High trademark	0.193 (0.369)	−0.000 (0.005)	0.006 (0.006)	−0.005 (0.007)	0.051 (0.051)
High IP × *High trademark*	−1.014** (0.425)		−0.018* (0.010)		−0.165 (0.154)
Sales growth	0.121 (0.352)	0.000 (0.002)	0.000 (0.002)	0.002 (0.003)	0.002 (0.003)
Entry rate	1.329 (0.911)	−0.019 (0.030)	−0.018 (0.030)	0.003 (0.032)	0.009 (0.032)
SAP penetration	0.576 (0.587)	0.016** (0.007)	0.015** (0.007)	0.002 (0.014)	−0.005 (0.017)
SAP product overlap	−0.247 (0.462)	0.012 (0.008)	0.011 (0.008)	0.010 (0.010)	0.003 (0.012)
Age	−0.046 (0.085)	0.003*** (0.001)	0.003*** (0.001)	0.002 (0.002)	0.001 (0.003)
Age^2	0.000 (0.002)	−0.000** (0.000)	−0.000** (0.000)	0.000 (0.000)	0.000 (0.000)
Log *Employee*	−0.172 (0.145)	0.004* (0.002)	0.004* (0.002)	−0.005 (0.005)	−0.004 (0.005)
Corporate investment	−0.098 (0.475)	−0.012*** (0.004)	−0.011** (0.004)	−0.022* (0.013)	−0.012 (0.011)
Private investment	0.119 (0.215)	−0.003* (0.001)	−0.003** (0.001)	0.007 (0.007)	0.004 (0.006)
VC investment	−0.953** (0.496)	0.002 (0.009)	0.002 (0.009)	−0.017 (0.017)	−0.015 (0.017)
Log *Publication*	1.529*** (0.478)	0.021 (0.016)	0.022 (0.016)	0.025* (0.015)	0.033* (0.018)
County employment	0.356** (0.171)	0.002* (0.001)	0.002 (0.001)	0.003 (0.002)	0.002 (0.002)
Year dummies	Yes	Yes	Yes	Yes	Yes
Overidentification test (p-value)				0.361	0.213
First stage F-statistic (p-value)				0.017	0.003 (*High IP*) 0.002 (Interaction)
Marginal effects	*dy/dx*	*dy/dx*	*dy/dx*	*dy/dx*	*dy/dx*
High IP (average)	0.019 (0.013)	0.017* (0.010)	0.023* (0.012)	0.288* (0.155)	0.289* (0.167)
High IP (*High trademark* = 0)	0.035* (0.019)		0.028** (0.014)		0.338* (0.203)
High IP (*High trademark* = 1)	0.004 (0.005)		0.010 (0.010)		0.172 (0.115)
High trademark (average)	−0.005 (0.005)	−0.000 (0.005)	0.000 (0.005)	−0.005 (0.007)	0.004 (0.009)
High trademark (*High IP* = 0)	0.002 (0.003)		0.006 (0.006)		0.051 (0.051)
High trademark (*High IP* = 1)	−0.029* (0.017)		−0.013 (0.009)		−0.114 (0.103)

Notes. An intercept is included in all specifications. Heteroskedasticity robust standard errors (clustered by firm) are in parentheses. Column (1) shows the results of a GEE model with probit link function. This model includes estimates for the time averages of the explanatory variables (not shown) used as controls for the unobserved firm fixed effects. Columns (2) and (3) use a fixed effects LPM. Columns (4) and (5) present the results from a fixed effects LPM using the instrumental variable method. The number of observations in all specifications is 6,498 (unbalanced panel of 1,220 firms observed over a nine-year period), except for models (4) and (5), where some observations are dropped due to insufficient within-group variance in the instrumental variables, with a resulting $N = 6{,}381$ (1,103 firms).

*$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.

5.2. Tests of Hypothesis 3

To examine whether the empirical evidence in our data is consistent with Hypothesis 3, we use the *Sales growth* of the ISV's target markets as a measure of profit potential and interact this variable with the ISV's IP protection and downstream capabilities. The results are presented in Table 5, which reports both the survival and probability models. In column (1) we present the results of the baseline Cox hazard model, where these interactions are added to the analysis. In column (2) we add the full set of control variables. As an alternative measure of market profit opportunities, we replace the *Sales growth* variable with product market *Entry rate* and run a similar Cox model, and we report the results in column (3). In addition, we estimate a discrete-time hazard model with complementary log–log link function, as well as a split population survival model, and present the results in columns (4) and (5), respectively. To account for unobserved firm heterogeneity, we present estimates from a fixed effects LPM in column (6). In column (7) we include estimates from a fixed effects LPM that instruments for *High IP* and its interaction with *Sales growth* using four instruments: a dummy for the regime change treatment group, the interaction of treatment group with a 1999 dummy, the interaction of treatment group with *Sales growth*, and the three-way interaction of treatment group, *Sales growth*, and the 1999 dummy. We have also experimented with controlling for the log of market sales; our results in these regressions are qualitatively similar to those without these controls.

We find that the effect of an ISV's IP protection on its decision to join a platform is significantly greater in high-growth markets. For example, the results of the baseline model with full set of controls (in column (2)) suggest that although, on average, the discrete change of *High IP* from 0 to 1 is associated an increase in the hazard of partnering of 141.6%, the effect is magnified (287.1% increase in partnering hazard, $p < 0.001$) if the sales growth of the ISV's target market is at the 90th percentile. On the other hand, the effect of *High IP* is diminished (9.9% increase in partnering hazard) and not statistically significant if market sales growth is at the 10th percentile. These results are robust to a variety of different stochastic assumptions and different ways of measuring market growth opportunities in columns (1)–(5). However, we are unable to identify a larger marginal effect for *High IP* in high-growth markets using only within-firm variation in our data (columns (6) and (7)); this reflects the relatively small number of firms switching to high IPR over our sample, making estimation of the interaction difficult. Moreover, we find no evidence that the marginal effect of trademarks is greater in rapidly growing markets. As mentioned previously, this is consistent with the idea that although downstream capabilities may be more valuable as a defense against the increased threat of platform owner entry associated with growing markets, they also tend to be relatively more effective when the application industries are relatively more mature.

6. Discussion and Conclusion

In this paper, we find evidence that an ISV's ownership of formal IPR such as patents and copyrights is associated with a significant increase in the likelihood of partnership with a leading enterprise software platform provider. Ownership of marketing or service capabilities is similarly associated with an increase in the tendency toward partnership. We further find that the presence of one appropriability mechanism weakens the marginal effect of the other on the likelihood and timing of partnership. Last, we provide evidence that the marginal effect of formal IPR is strongest in high-growth markets.

In our setting the decision to join a platform is driven by a clear trade-off. Joining the platform lowers users' expected sunk costs of integrating the ISV's applications with the platform and signals compatibility with a range of complementary economic activities specific to the platform. In this way, joining the platform increases the net benefits for an existing platform user to adopt the ISV's applications and therefore makes it easier for the ISV to sell to the platform owner's installed base. However, platform partnership may increase the risk of platform owner entry into the complementary market and a subsequent profit squeeze. By providing evidence that ownership of appropriability mechanisms increases the likelihood of platform partnership, our results suggest that such mechanisms increase the excludability of ISV innovation and stimulate incentives to provide complementary applications for the platform.

Our results highlight the role that appropriability mechanisms play in ameliorating a fundamental problem in platform governance: a platform owner's inability to commit not to squeezing providers of complementary products and services. The role of such mechanisms has thus far received little attention in the platform literature, possibly because of the widespread skepticism on the role of patents in protecting innovation in the software industry: alternative appropriation strategies such as secrecy are often considered a far more effective alternative. We interpret our findings as reflecting the effectiveness of formal IPR as an appropriability mechanism relative to alternatives. After all, in our setting, cooperation takes the form of partnerships that ensure software compatibility with the platform, a delicate process that leads to the risk of disclosing software design information

Table 5 The Role of IPR in Growing Markets

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables	Cox baseline	Cox with controls	Cox with controls	Cloglog	Split pop. cure model	LPM with FE	LPM with FE and IV
High IP	−6.110***	−6.429***	0.832	−5.584**	−6.087**	0.016	0.154
	(2.191)	(2.480)	(0.579)	(2.616)	(3.006)	(0.016)	(0.150)
High trademark	3.352***	2.599**	1.008**	1.987*	2.299	0.016*	0.013
	(1.206)	(1.293)	(0.494)	(1.138)	(1.616)	(0.009)	(0.028)
High IP × *High trademark*	−0.710	−1.031	−1.309*	−0.979	−1.122	−0.018*	−0.014
	(0.728)	(0.762)	(0.739)	(0.822)	(0.762)	(0.010)	(0.023)
Sales growth	−4.986***	−5.573***	−0.788	−4.447**	−5.285**	0.000	−0.014
	(1.794)	(2.043)	(1.377)	(2.222)	(2.580)	(0.002)	(0.017)
Sales growth × *High IP*	6.123***	6.463***		5.766***	6.213**	0.010	0.092
	(1.772)	(2.050)		(2.200)	(2.579)	(0.010)	(0.107)
Sales growth × *High trademark*	−1.494	−1.357		−1.015	−1.177	−0.008	−0.014
	(0.995)	(1.104)		(0.940)	(1.377)	(0.006)	(0.023)
Entry rate		2.277*	−3.572	3.920***	2.991**	−0.017	0.004
		(1.350)	(3.087)	(1.098)	(1.498)	(0.030)	(0.032)
Entry rate × *High IP*			7.128**				
			(3.101)				
Entry rate × *High trademark*			1.041				
			(3.630)				
SAP penetration		2.067**	2.133**	2.124**	1.551	0.014**	−0.000
		(0.818)	(0.904)	(0.841)	(1.094)	(0.007)	(0.014)
SAP product overlap		−0.814	−0.875	−0.931*	−0.811	0.011	0.012
		(0.534)	(0.535)	(0.536)	(0.552)	(0.008)	(0.010)
Age		−0.030	−0.049	−0.100	−0.052	0.003***	0.002
		(0.124)	(0.125)	(0.120)	(0.123)	(0.001)	(0.002)
Age^2		0.000	0.001	0.002	0.000	−0.000**	0.000
		(0.005)	(0.006)	(0.005)	(0.005)	(0.000)	(0.000)
Log *Employee*		0.720***	0.703***	0.725***	0.714***	0.004*	−0.004
		(0.151)	(0.142)	(0.161)	(0.162)	(0.002)	(0.005)
Corporate investment		0.347	0.290	0.206	0.234	−0.012***	−0.022*
		(0.653)	(0.629)	(0.704)	(0.623)	(0.004)	(0.013)
Private investment		−0.512	−0.475	−0.590	−0.557	−0.003*	0.009
		(0.432)	(0.436)	(0.452)	(0.383)	(0.001)	(0.008)
VC investment		0.780*	0.858**	0.794*	0.773**	0.002	−0.013
		(0.406)	(0.394)	(0.422)	(0.393)	(0.009)	(0.016)
Log *Publication*		0.009	−0.044	−0.004	0.019	0.022	0.024
		(0.250)	(0.255)	(0.233)	(0.346)	(0.016)	(0.015)
County employment		0.174	0.160	0.191	0.158	0.002	0.003
		(0.183)	(0.176)	(0.191)	(0.185)	(0.001)	(0.002)
Marginal effects	*ey/dx*	*ey/dx*	*ey/dx*	*ey/dx*	*ey/dx*	*dy/dx*	*dy/dx*
High IP (*Sales growth/Entry rate* = average)	1.608***	1.416***	0.746*	1.395**	1.416***	0.023*	0.270*
	(0.581)	(0.499)	(0.427)	(0.552)	(0.541)	(0.012)	(0.061)
High IP (*Sales growth* = 10%)	0.361	0.099		0.223	0.150	0.021*	0.251*
	(0.618)	(0.500)		(0.520)	(0.494)	(0.012)	(0.136)
High IP (*Sales growth* = 90%)	2.987***	2.871***		2.691***	2.815***	0.025**	0.291*
	(0.769)	(0.808)		(0.901)	(0.995)	(0.013)	(0.156)
High IP (*Entry rate* = 10%)			0.284				
			(0.485)				
High IP (*Entry rate* = 90%)			1.380***				
			(0.492)				

Notes. Columns (1) and (2) show the results of Cox proportional hazard models, column (3) uses a complementary log–log regression, and column (4) uses a split sample population cure model. Column (5) uses a fixed effects LPM, and column (6) uses a LPM with instrumental variables for *High IP* and *Sales growth* × *High IP*. Marginal effects in columns (1)–(5) are elasticities (and so are denoted by the notation *ey/dx*), whereas those in columns (6) and (7) are derivatives (and denoted by *dy/dx*). Number of observations: 6,498 (1,220 firms). Heteroskedasticity robust standard errors (clustered over firms) are in parentheses, except in column (4), which uses classical (independent and identically distributed) standard errors.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

to the platform owner, and therefore lowers the effectiveness of popular alternative protection mechanisms such as trade secrecy.

Our results have important implications for where platforms are most likely to grow. Platforms will be less likely to grow in settings with little formal means of IP protection and, in particular, where patent and copyright protection is weak. They will be relatively more successful when ISVs are more effectively able to secure returns from their innovations through patents, copyrights, and downstream capabilities. Under such conditions, ISVs are more likely to enter into markets complementary to the platform and produce platform-compatible applications. Such entry will enhance the platform's value and expedite its adoption, setting into motion a virtuous cycle of indirect network effects.

Relatedly, our results have implications for platform owners and policy makers seeking to encourage entry in complementary markets. As noted earlier, platform owners have employed a variety of alternative mechanisms to commit not to squeezing potential entrants. Our results suggest that such mechanisms may be less important in settings where ISVs can use IPR to appropriate the returns from their inventions. Furthermore, policy makers occasionally take regulatory actions when the potential for a profit squeeze leads to levels of entry that are too low from a social welfare perspective. For example, recent enforcement actions against Microsoft by antitrust regulators in the United States and European Union can be viewed as an attempt to encourage entry in complementary markets (Miller 2008). Our results suggest that such policies may be less necessary in settings with strong IPR, and that accurate understanding of the appropriability environment in which a market is situated can be usefully employed to gauge appropriate policy responses.

Finally, our research makes a contribution to the markets for technology literature by investigating the role of appropriability in shaping the commercialization strategies of start-ups (Arora and Ceccagnoli 2006, Arora et al. 2001, Dechenaux et al. 2008, Gans et al. 2002). Our study is based on a different approach for studying the implications of appropriability than most research in that literature. In contrast to most prior work that has utilized cross-industry survey data to test hypotheses, our approach uses secondary data collected from a single industry. This approach is appealing because it allows the researcher to extend and advance well-recognized general frameworks (e.g., Gans and Stern 2003) to accommodate idiosyncratic industry conditions, furthering our knowledge of markets for technology. We hope that our research will encourage additional work that applies this approach to other contexts.

Acknowledgments

The authors thank Kevin Boudreau, Erik Brynjolfsson, Annabelle Gawer, Shane Greenstein, Paul Hofmann, Marco Iansiti, Karim Lakhani, Kristina McElheran, Robert Seamans, Sandra Slaughter, Scott Stern, Marie Thursby, Marshall Van Alstyne, Feng Zhu, and seminar participants at Carnegie Mellon University, Duke University, Harvard Business School, the Wharton School, University of Maryland, University College London, the 30th International Conference on Information Systems, the 9th Annual Roundtable for Engineering Entrepreneurship Research, the Kauffman Entrepreneurship Mentoring Workshop (2010), the Academy of Management Symposium on Competitive Dynamics in Platform-Based Markets (2010), and the Industry Studies Association Conference (2010) for helpful comments and suggestions. They also thank Gérard Cachon (the department editor), the associate editor, and the reviewers for their constructive comments and insights, and the many executives of enterprise software companies who shared their thoughts with them. They acknowledge funding from an NET (Networks, Electronic Commerce, and Telecommunications) Institute summer grant (2009). Peng Huang thanks the Kauffman Foundation for providing funding for this research through a Kauffman Dissertation Fellowship. Chris Forman acknowledges funding from the Alfred P. Sloan Foundation through an Industry Studies Fellowship. Chris Forman and D. J. Wu acknowledge a Kauffman Foundation/Georgia Research Alliance Grant for the Study of Entrepreneurship and Productivity. All opinions and errors are the authors' own.

References

Ai C, Norton EC (2003) Interaction terms in logit and probit models. *Economic. Lett.* 80(1):123–129.

Angrist JD, Pischke J-S (2009) *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press, Princeton, NJ).

Arora A, Ceccagnoli M (2006) Patent protection, complementary assets, and firms' incentives for technology licensing. *Management Sci.* 52(2):293–308.

Arora A, Fosfuri A, Gambardella A (2001) *Markets for Technology: The Economics of Innovation and Corporate Strategy* (MIT Press, Cambridge, MA).

Bader MA (2006) *Intellectual Property Management in R&D Collaborations: The Case of the Service Industry Sector* (Physica-Verlag, Heidelberg, Germany).

Becchetti L, Paganetto L (2001) The determinants of suboptimal technological development in the system company-component producers relationship. *Internat. J. Indust. Organ.* 19(9):1407–1421.

Bessen J, Hunt RM (2007) An empirical look at software patents. *J. Econom. Management Strategy* 16(1):157–189.

Boudreau K (2010) Open platform strategies and innovation: Granting access vs. devolving control. *Management Sci.* 56(10): 1849–1872.

Ceccagnoli M, Rothaermel FT (2008) Appropriating the returns from innovation. Libecap GD, Thursby MC, eds. *Technological Innovation: Generating Economic Results* (Elsevier, Oxford, UK), 11–34.

Ceccagnoli M, Forman C, Huang P, Wu DJ (2012) Co-creation of value in a platform ecosystem: The case of enterprise software. *MIS Quart.* 36(1):263–290.

Chamberlain G (1980) Analysis of covariance with qualitative data. *Rev. Econom. Stud.* 47(1):225–238.

Cockburn IM, MacGarvie MJ (2009) Patents, thickets and the financing of early-stage firms: Evidence from the software industry. *J. Econom. Management Strategy* 18(3):729–773.

Cockburn IM, MacGarvie MJ (2011) Entry and patenting in the software industry. *Management Sci.* 57(5):915–933.

Cohen WM, Nelson RR, Walsh JP (2000) Protecting their intellectual assets: Appropriability conditions and why U.S. manufacturing firms patent (or not). NBER Working Paper 7552, National Bureau of Economic Research, Cambridge, MA.

Colombo MG, Grilli L, Piva E (2006) In search of complementary assets: The determinants of alliance formation of high-tech start-ups. *Res. Policy* 35(8):1166–1199.

Dechenaux E, Goldfarb B, Shane S, Thursby M (2008) Appropriability and commercialization: Evidence from MIT inventions. *Management Sci.* 54(5):893–906.

Economides N (1988) The economics of trademarks. *Trademark Reporter* 78:523–539.

Evans DS, Hagiu A, Schmalensee R (2006) *Invisible Engines: How Software Platforms Drive Innovation and Transform Industries* (MIT Press, Cambridge, MA).

Farrell J, Katz ML (2000) Innovation, rent extraction, and integration in systems markets. *J. Indust. Econom.* 48(4):413–432.

Fosfuri A, Giarratana MS, Luzzi A (2008) The penguin has entered the building: The commercialization of open source software products. *Organ. Sci.* 19(2):292–305.

Gambardella A, Giarratana MS (2006) Innovations for products, innovations for licensing: Patents and downstream assets in the software security industry. Working paper, Bocconi University, Milan, Italy. http://ssrn.com/abstract=935210.

Gans JS, Stern S (2003) The product market and the market for "ideas": Commercialization strategies for technology entrepreneurs. *Res. Policy* 32(2):333–350.

Gans JS, Hsu DH, Stern S (2002) When does start-up innovation spur the gale of creative destruction? *RAND J. Econom.* 33(4):571–586.

Gawer A (2009) *Platforms, Markets, and Innovation* (Edward Elgar, Cheltenham, UK).

Gawer A, Cusumano MA (2002) *Platform Leadership: How Intel, Microsoft, and Cisco Drive Industry Innovation* (Harvard Business School Press, Boston).

Gawer A, Henderson R (2007) Platform owner entry and innovation in complementary markets: Evidence from Intel. *J. Econom. Management Strategy* 16(1):1–34.

Graham S, Mowery D (2003) Intellectual property protection in the U.S. software industry. Cohen WM, Merrill SA, eds. *Patents in the Knowledge-Based Economy* (National Academies Press, Washington, DC), 219–258.

Graham S, Mowery D (2006) Intellectual property rights in the software industry: Implications for open innovation. Chesbrough H, Vanhervebeke W, West J, eds. *Open Innovation: Researching a New Paradigm* (Oxford University Press, Oxford, UK), 184–204.

Graham SJH, Merges RP, Samuelson P, Sichelman T (2010) High technology entrepreneurs and the patent system: Results of the 2008 Berkeley patent survey. *Berkeley Tech. Law J.* 24(4): 1255–1328.

Hall BH (2009) Business and financial method patents, innovation, and policy. *Scottish J. Political Econom.* 56(4):443–473.

Hall BH, Helmers C (2011) Innovation and diffusion of clean/green technology: Can patent commons help? NBER Working Paper 16920, National Bureau of Economic Research, Cambridge, MA.

Hall BH, MacGarvie M (2010) The private value of software patents. *Res. Policy* 39(7):994–1009.

Hall BH, Jaffe AB, Trajtenberg M (2002) The NBER patent citations data file: Lessons, insights and methodological tools. Jaffe AB, Trajtenberg M, eds. *Patents, Citations and Innovations: A Window on the Knowledge Economy* (MIT Press, Cambridge, MA), 403–459.

Heeb R (2003) Innovation and vertical integration in complementary markets. *J. Econom. Management Strategy* 12(3):387–417.

Hitt LM, Wu DJ, Zhou X (2002) Investment in enterprise resource planning: Business impact and productivity measures. *J. Management Inform. Systems* 19(1):71–98.

Hosmer DW, Lemeshow S (1999) *Applied Survival Analysis: Regression Modeling of Time to Event Data* (Wiley-Interscience, New York).

Katz ML, Shapiro C (1986) Technology adoption in the presence of network externalities. *J. Political Econom.* 94(4):822–841.

Klepper S (1996) Entry, exit, growth, and innovation over the product life cycle. *Amer. Econom. Rev.* 86(3):562–583.

Lerner J (2002) Where does State Street lead? A first look at finance patents, 1971 to 2000. *J. Finance* 57(2):901–930.

Lerner J, Zhu F (2007) What is the impact of software patent shifts? Evidence from *Lotus v. Borland*. *Internat. J. Indust. Organ.* 25(3):511–529.

Mann RJ, Sager TW (2007) Patents, venture capital, and software start-ups. *Res. Policy* 36(2):193–208.

Miller AR, Tucker C (2009) Privacy protection and technology diffusion: The case of electronic medical records. *Management Sci.* 55(7):1077–1093.

Miller DA (2008) Invention under uncertainty and the threat of ex post entry. *Eur. Econom. Rev.* 52(3):387–412.

Mundlak Y (1978) On the pooling of time series and cross section data. *Econometrica* 46(1):69–85.

Nahm J (2004) Open architecture and R&D incentives. *J. Indust. Econom.* 52(4):547–568.

Pang A (2007) SAP Ecosystem becomes a booming economy. Research report, International Data Corporation (IDC), Framingham, MA.

Parker G, Van Alstyne MW (2012) Innovation, openness, and platform control. Working paper, Tulane University, New Orleans. http://ssrn.com/abstract=1079712.

Petersen MA, Rajan RG (1994) The benefits of lending relationships: Evidence from small business data. *J. Finance* 49(1):3–37.

Prentice PL, Gloeckler LA (1978) Regression analysis of grouped survival data with applications to breast cancer data. *Biometrics* 34(1):57–67.

Puranam P, Singh H, Zollo M (2006) Organizing for innovation: Managing the coordination-autonomy dilemma in technology acquisitions. *Acad. Management J.* 49(2):263–280.

Rysman M, Simcoe T (2008) Patents and the performance of voluntary standard-setting organizations. *Management Sci.* 54(11):1920–1934.

SAP (2005a) Powered by SAP Netweaver Partner Program FAQ. Accessed September 8, 2012, http://www.sdn.sap.com/irj/scn/index?rid=/library/uuid/e663625c-0c01-0010-6e8d-9ec8c49d1ffc.

SAP (2005b) SAP annual report 2005. Accessed September 8, 2012, http://www.sap.com/corporate-en/investors/reports/annualreport/2005/pdf/2005_SAP_Annual_Report.pdf.

SAP (2009) SAP facts and figures. Accessed April 21, 2009, http://www.sap.com/about/investor/pdf/SAP_FactSheet.pdf.

Schmidt P, Witte AD (1989) Predicting criminal recidivism using "split population" survival time models. *J. Econometrics* 40(1):141–159.

Shapiro NL (2005) Memorandum and order, *AMC Technology LLC v. SAP AG*, November 3. United States District Court for the Eastern District of Pennsylvania. Accessed September 8, 2012, http://www.paed.uscourts.gov/documents/opinions/05D1367P.pdf.

Teece DJ (1986) Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Res. Policy* 15(6):285–305.

Utterback JM, Abernathy WJ (1975) A dynamic model of product and process innovation. *Omega* 3(6):639–656.

Wen W, Ceccagnoli M, Forman C (2011) Patent pools, thickets, and open source entry by start-up firms. Working paper, Georgia Institute of Technology, Atlanta.

Williamson OE (1991) Comparative economic organization: The analysis of discrete structural alternatives. *Admin. Sci. Quart.* 36(2):269–296.

Wooldridge JM (2001) *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge, MA).

Zeger SL, Liang K-Y, Albert PS (1988) Models for longitudinal data: A generalized estimating equation approach. *Biometrics* 44(4):1049–1060.