Smartphones, Apps, and Digital Flows: Platform Competition in the Mobile Industry

By

BRYAN PON

B.A. (University of California, Santa Barbara) 1999
M.S. (University of California, Davis) 2012
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Geography

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

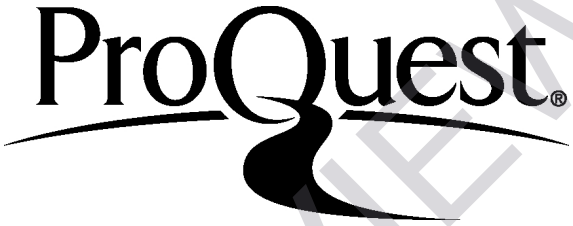
Martin Kenney, chair

Chris Benner

Andrew Hargadon

Committee in charge

2015



Abstract

The mobile telecommunications industry is in the process of a dramatic transformation into the smartphone industry, as new firms from the computer and internet sectors have used new technologies and business models to displace the incumbents. The key organizing structure of this new smartphone industry is the software "platform," a model which defined the PC industry for decades but is also present in other industries, such as console video games. Software platforms have distinct economic properties that shape competitive strategy, including the presence of positively reinforcing network effects, which lead to increasing returns to scale and the potential for winner-take-all markets. In the smartphone industry, the platforms by Apple, Google, Microsoft, Amazon, and Xiaomi also function as two-sided markets, bringing together distinct groups (e.g., app developers and end-users) who transact through the platform. In the work that follows, I explore how these and other platform dynamics are exploited by the key firms in their apparent strategies, and the implications for competition in the industry. The analysis shows that new organizational forms such as hybrid platforms, where firms such as Amazon and Xiaomi have appropriated open-source Android for their own proprietary platforms, and nested platforms, where firms such as Line and Facebook build distinct platforms within other platforms, challenge the traditional platform model and our understanding of how firms exercise platform control. I argue that the "bottleneck" or control point of smartphone platforms has moved away from the operating system, and up the stack to cloud-based services. The final part of the analysis studies the app economy, and how the platform-mediated "app stores" are shaping participation and value capture. Using spatial analysis, I map the geography of digital flows of apps between developers and the most lucrative markets, revealing clear patterns of inter-regional trade and insular domestic markets.

Acknowledgements

I would like to thank my advisor, Martin Kenney, for his guidance, insights, and support in completing this work. He broadened my understanding of academic publishing, and academia in general, and I remain extremely grateful for all of his time and mentorship. Also, heartfelt appreciation and respect for those in my masters and doctoral cohort, especially Scoot, Magdalena, Pascoe, Mariana, and Crandall (and Brad who is basically an honorary member). You guys have challenged me from our first class together, and you made grad school so much more than coursework and research. I hold out hope that someday our frisbee or innertube water polo team will rise again….We may be old, but at least we're slow. And of course, immense gratitude and love to my wife Erica, who not only brought home (and often cooked) the bacon for the last four years, but broadened my perspective, checked my cynicism, and offered the moral support I often needed to finish this work.

Foreword

In lieu of a traditional, book-length dissertation, this dissertation comprises four academic articles on smartphone platforms, three of which have been published in peer-reviewed journals, and one that is in the process of being submitted.

In the following section, Part I, I introduce the relevant theoretical constructs and bodies of literature that my research is based in, and situate the work within these. This section provides a high-level discussion of all the themes used in the four articles, summarized and put into context with one another. Part II, III, IV, and V are each comprised of an academic article covering different facets of my research on smartphone platforms and the app economy. Part VI integrates many of these themes into a broader discussion with ideas for further research. This is followed by a complete bibliography that contains references not listed in the individual articles.

The three published articles, parts II-IV, can be found online as the following:

- Kenney, Martin, and Bryan Pon. "Structuring the smartphone industry: is the mobile internet OS platform the key?" Journal of Industry, Competition and Trade 11.3 (2011): 239-261.
- Pon, Bryan, Timo Seppälä, and Martin Kenney. "Android and the demise of operating system-based power: Firm strategy and platform control in the post-PC world." Telecommunications Policy 38.11 (2014): 979-991.
- Pon, Bryan, Timo Seppälä, and Martin Kenney. "One Ring to Unite Them All: Convergence, the Smartphone, and the Cloud." Journal of Industry, Competition and Trade (2015) 15:21–33

Part I. Introduction

Overview

New developments in information and communications technology are continuously shaping—and being shaped by—our socioeconomic structures, behaviors, and systems. By enhancing our ability to create, process, store, and communicate symbols and meaning, these technologies can have deep impacts on how we organize and coordinate activity at multiple scales. While this has always been the case, the growing significance of information in the global economy gives these technologies even more transformative power. Because while information and knowledge have always been a critical input into new innovations, we have entered an age of "informationalism," where information and knowledge are no longer simply inputs in the production function, but the innovative products and services themselves (Castells, 1996: 31). This productive shift is happening concurrently—and precipitating in many ways—the broad set of economic reconfigurations we term globalization, leading to ever more complex interconnections between firms, states, and public-sector actors across both political and industry boundaries. The significance of these changes is immense, and they offer both great opportunity and peril. For example, there is the imminent potential for providing computing technology and internet access to billions of people for whom even 10 years ago the idea would have seemed naïve. Yet there is real danger in the winner-take-all nature of networked ICT industries, as only a handful of private firms—all based in the United States—seem to be in control over vast swathes of this new technological landscape. And with a device so personal that it can track our movements, communications, even our health, the security and privacy risks of mobile technology are difficult to understate.

Against this backdrop of dynamic change and complexity, social scientists and economic geographers especially face compelling opportunities, for the toolkit of geography is uniquely well-suited to interpreting and making sense of these technological and economic transformations. The scale and degree of these changes—enhanced by seemingly endless hype—may require re-examination of long-held beliefs, but fundamental principles are unlikely to change quickly, if at all. We know that distance is still alive, people collaborate better in person, and cities still matter (in fact

increasingly so). The task for geographers, then, is to develop nuanced understandings of how these changes are impacting the outcomes we care about, and offer models and insights that can help policymakers and practitioners interpret the complexity in meaningful ways.

The smartphone era

The smartphone industry of 2015 is situated within an increasingly complex technological and economic space, characterized by multi-sided markets, network effects, a variety of business models, ever-advancing technologies, and the layered nature of the computer-telephony industry itself. From their roots in other industries, many of the key firms—especially Apple, Google, and Microsoft—have different core competencies and business models, leading to multi-dimensional and asymmetric competition. Across this highly complex and dynamic market environment, the dominant paradigm shaping firm strategy, industry structure, and economic outcomes is that of a software "platform," which can be understood as the organization of interconnected technologies and processes which enable and incentivize external firms to build and create shared value, which due to network effects can see increasing returns to scale and the potential for winner-take-all markets.

The importance of software platforms isn't new. From Microsoft Windows to Salesforce.com to Facebook to Nintendo, software platforms often define information and communication technology industries, spawn new markets, and generally influence technological development in these sectors (Evans et. al, 2006). Scholars and management consultants have identified platform control as a key feature for business success in the ICT industries. Michael Cusumano, drawing upon his studies of Microsoft, Cisco, and Intel, concluded that the winners of technological competitions are often "who has the best platform strategy and the best ecosystem to back it up (Cusumano 2010)."

While much of the early research on platforms explored the personal computer (PC) industry (Bresnahan and Greenstein, 1999; Cusumano and Yoffie, 1998; West, 2003) and the historical development of its iconic platform firms, platform scholars have in recent years shifted attention to the mobile industry (e.g., Gonçalves and Ballon, 2010; Basole and Karla, 2012; Heitkoetter et. al, 2012; West 2010), which entered a new phase—i.e., the smartphone era—with the launch of the

Apple iPhone in 2007. The computing capability of the iPhone was quickly recognized by developers, who clamored for access, and in 2008 Apple agreed to let 3rd-party developers write software, "apps," that could be installed on the iPhone. That same year, Google released the first smartphone on its competing operating system, Android, also opening it up to 3rd-party developers, and both platforms saw a dramatic and rapid uptake in the number of apps being built and downloaded. Smartphone use quickly became centered on task-specific apps, developed primarily by independent 3rd-parties (though some key apps, for example Google Maps, were built by the platform owners) and sold or distributed through the respective online virtual marketplaces, giving birth to the "app economy." Both iOS and Android have seen dramatic growth in worldwide adoption, with 16% and 80% of the market, respectively, and Android has become the fastest-growing technology platform in history, reaching the milestone of 1 billion active users in just 5 years, faster than even Facebook (almost 9 years).[1]

Using the platform lens on the smartphone industry is valuable both because it is fundamental to the structure, operation, and power dynamics of the industry, but also because the very nature of a platform—the theoretical construct—seems to be under challenge and evolving as a result of current firm strategies and business models (Pon et. al, 2014). Because technology platforms are becoming more important in the global economy—firms such as Uber, AirBnB, oDesk/Elance, and others are restructuring value chains in many sectors—this latter aspect can lead to more generalizable insights that cross industry boundaries and inform our understanding of these dynamics more broadly.

This inquiry has explored a few specific platform themes across the four journal articles. The first theme involves basic strategy for platform firms, who must determine issues such as how open to make the platform, how to set prices for participation, and how to capture value. In Kenney and Pon (2011), I compare the conventional wisdom for platform firms (Gawer and Cusumano, 2002) against apparent strategies from Apple, Google, Microsoft, and Nokia, who all have different core competencies that shape their business models. I found that the open platform strategy of Google, and

[1] ASYMCO, http://www.asymco.com/2013/01/16/the-race-to-a-billion-2012-update/; http://www.engadget.com/2014/06/25/google-io-2014-by-the-numbers/

its radically different monetization model (advertising), makes its strategy hard to reconcile with conventional platform theory.

Building on this theme in Pon et. al (2014), I then explore how firms can compete and differentiate if the fundamental technology of the platform—the operating system—is the same. Through case studies of Google, Amazon, and Xiaomi, who have all built proprietary platform ecosystems on top of an open-source operating system, I evaluate each firm's ability to serve specific "gatekeeper roles" (Ballon, 2009) that can serve as control points within the ecosystem.

The theme of platform control recurs in Pon et. al (2015), where I more explicitly explore how the shifting landscape of the smartphone industry, and especially the asymmetric competition by Google, has led to new "bottlenecks" or locations for control. I develop the hypothesis that the operating system has ceased to be the most important control point, and that instead, it is the virtual markets and internet-enabled services (via apps) that are the most important for user lock-in and platform control.

The spatial and structural importance of the app stores as virtual markets is the focus for the final paper (Pon et. al, forthcoming). In this work, I use location data on app developers and the national markets where they distribute in order to trace "digital flows" of data and money over the telecommunications network and across borders. This geographical analysis of how individual firms and nations are able to participate and capture value from the growing app economy paints a global picture of productive capacity in the digital economy.

The rise of platforms

The growing literature on technology "platforms" was popularized by Michael Cusumano in 1998, when he began applying concepts from the technology standards literature to develop business management strategies for firms in high technology industries (Cusumano and Yoffie, 1998; Cusumano and Gawer, 2001). Since that time, his former student, Annabelle Gawer, has led many research efforts to explore platform dynamics and develop agreement on common definitions (especially Gawer, 2010; Gawer 2011; Gawer 2014). This section provides a brief overview of

platform characteristics and a typology of platforms, followed in the next section by more detailed descriptions of the antecedents of this emerging platform theory.

The term "platform" has been used across multiple fields, but the most influential antecedents are from systems engineering and industrial economics. In the former, platforms have been conceived of as modular structures that use technology standards to support flexible, iterative development. Baldwin and Woodard (2009) lay out a structural definition of a product platform as comprising three elements: a core technology that serves as a foundation, a range of components that can connect to and extend this core, and the interfaces in-between. In the physical realm this would include, for example, a car chassis that is used across multiple models of car, where the core is the chassis, the components are suspension parts or body panels, and the interfaces are the specific dimensions and configurations of mounting brackets. For operating systems such as Android or iOS, the core technology is the operating system (OS), the components are apps and software services that are run on the OS, and the interfaces are the application programming interfaces (APIs) that define these connections between apps and the OS.

The other field that has adopted platforms as a construct is industrial economics, where a platform is a two-sided or multi-sided market for distinct user groups (e.g., "buyers" and "sellers") to conduct transactions. This view is founded on the idea of network effects, where the value of participating for one user group is directly tied to the participation of the other user group (David, 1985; Katz and Shaprio, 1985; 1994). Examples of two-side markets include credit cards, shopping malls, and singles clubs; in the former, for example, the value of having a Visa or American Express card is proportional to the number of merchants who will accept it (and vice versa). Economists studying two-sided markets have tended to focus on pricing as a lever to incentivize participation and establish governance strategies (Rochet and Tirole, 2003; Eisenmann et. al, 2006). For example, one strand of this research seeks to determine how to bring both sides of the platform "on board" simultaneously, which usually requires some kind of subsidy as an incentive (Evans, 2003).

Because the broad descriptions above include everything from a car chassis to credit cards, it is important to distinguish a subset of platforms based on software, such as Microsoft Windows, Sony PlayStation, Linux, Apple Mac and iOS, and Google Android. While many, if not most, of today's

platforms may be mediated by software (e.g., car chassis designed with CAD), software platforms are distinct in that the core technology, the components that connect to the core, and the interfaces in-between are all comprised of software. Of course, as evident by these examples, even pure software platforms require hardware to function, so the PC, gaming console, and smartphone devices are all integral elements of these software platforms (Evans et. al, 2006). Yet the nature of software, as a digitized information good, creates special conditions for platforms with software as the key modular component. I describe the implications of digital goods in greater detail in the section "Markets in (digital) space."

The role of standards

The IT sector is predicated on, and built on, standards—codified specifications that determine how different components may connect and function together (Katz and Shapiro, 1985; Farrell and Saloner, 1985). Computing and communication technologies especially must adhere to standards at multiple levels, in both the software and hardware, in order to achieve interoperability, and hence, utility (Farrell and Saloner, 1992). Standards can be "de jure" and established through a formal process involving multiple stakeholders and overseen by a public agency or standards setting organization, or "de facto," when the technology becomes a standard simply by virtue of its degree of adoption or importance in the market (Garud et. al, 2002; Funk and Methe, 2001).[2] Of course, these two descriptors are at opposite ends of a spectrum, and many standards and their respective processes will exist at some point on the continuum between them (West, 2007). The often complex and contested processes by which standards are established reflects their potential significance in determining technological and economic outcomes, especially in sectors where network effects can lead to winner-take-all markets.

Much of the standards literature grew out of the observation that in many high technology industries, a single technology will become dominant, even if—in many cases—it is technologically inferior to alternatives (Rosenbloom and Cusumano 1985; Von Burg, 2001). Researchers have

[2] David and Greenstein (1990) further categorize into "unsponsored" and "sponsored" forms of de facto standards, with the former being those that have no clear backer but nevertheless exist in the public domain

suggested two key concepts in explaining these cases: first, there exists an element of "path-dependence" affecting the evolution and adoption of technologies, and this path dependence can be shaped by small events that over time drive large change[3] (David, 1985; Arthur, 1994). Secondly, in technology sectors defined by networks, there are self-reinforcing cycles that can produce winner-take-all markets.[4] This is because network technologies exhibit positive consumption externalities, whereby the value of being on the network increases for each additional user of the network. These "network effects" include direct network effects, where the value for an end-user increases when other end-users are on the same network (for example, the more people who use email, the more useful it is for each user). This is contrasted with indirect network effects, where the value of being on the network increases when there are more complementors providing products and services around the network (for example, the more email software available, the more attractive email is for end-users, and vice versa) (Katz and Shapiro, 1985; Shapiro and Varian, 1999). This positively reinforcing cycle of network effects leads to a fundamental shift in the economic model for network industries, where instead of the equilibrium-seeking effects of decreasing returns to scale, networks foster increasing returns and the potential for winner-take-all-markets (Arthur, 1989; 1994).

Systems design incorporates modularity

Building on the standards literature, researchers in the field of systems or engineering design have explored how modularity of systems—the ability of components to be separated and recombined in new ways—can lead to competitive advantage. By deciding to decouple components of a system in order to make it more modular, firms benefit from increased system flexibility, economies of scale from strategic re-use of core elements, faster iteration on new designs due to reduced integration

[3] The impact of path dependence has been contested. For example, Kay (2013) argues that the QWERTY design was sufficiently superior to Dvorak and others such that even with different historical events, QWERTY would always "win" the standards war. Liebowitz and Margolis (1995) define a typology of path dependence, and argue that the most extreme type—where "insignificant" events create lock-in to an inferior standard—is not supported theoretically.

[4] The classic business school case studies of the QWERTY keyboard (David, 1985), alternating current and direct current (David, 1988), or VHS and Betamax (Rosenbloom and Cusumano, 2012; Liebowitz and Margolis, 1995) demonstrate the potential for standards to define a winner-take-all market.

complexity, and (with industry-wide platforms) the potential for external innovation due to codified mechanisms for connecting to the platform (Garud and Kumaraswamy, 1995; Schilling, 2000; Gawer, 2010). The act of decoupling allows some components to evolve and change over time without affecting the function of other key components. In dynamic, quickly changing industries, this enables innovation to occur where needed, without requiring complete overhaul of the entire product each time a small change is made (Tiwana et. al, 2010). A critical element of product platforms is the specification of interfaces, which serve as the common language spoken between disparate components even if the components themselves have changed substantially (Schilling, 2000). At the software level, the interface typically consists of a set of defined communication protocols called application programming interfaces (APIs), which are what allow one software program (e.g., a smartphone app) to talk with another software program (e.g., the smartphone operating system). Finally, product platforms can exist at multiple scales: internal to the organization, throughout a supply chain, and across entire industries; the key distinction is the potential population of complementors who can innovate on the platform, ranging from only employees of the firm, to partner firms, to any number of firms or individuals who may or may not have contractual agreements with the lead platform firm (Gawer, 2010; 2014).

Technology platforms are therefore the result when a modular system, based on a collection of standards, is opened up to other firms so that they can develop components that connect to the core technology, and thereby participate in the value chain of the platform. The modularity allows external innovation and value creation from complementors, and enables a much quicker pace of innovation through the re-use of components. Of course, attracting and incentivizing participation in the platform in the first place often requires "winning" a standards battle with competing technologies, either through formal (e.g., de jure) or informal (de facto) standards-setting processes. And for those platforms in network industries, participation is driven by self-reinforcing network effects and increasing returns to scale, often leading to the establishment of a single dominant technology in a winner-take-all market.

Platform geographies

With its origins in technology standards, engineering design, and industrial economics, the platform literature has no appreciable spatial dimension. Given platforms' fundamental role in coordinating networks of firms and technologies across time and space, this lack of consideration for place is a significant gap in platform theory. Research from other social sciences, especially economic geography, can provide some guidance for how we may incorporate concepts of space into the platform construct, as there is a rich and growing body of literature that examines how firms and industries organize and coordinate economic activity (Coe et. al, 2008).

One of the most comprehensive models is the concept of a global production network (GPN), which although relatively recent, has roots in a number of different disciplines (Hess and Yeung, 2006). One prominent precursor is the "value chain" work by Michael Porter (1985; 1990) in the field of strategic management. By analyzing firm activities at a more granular level, Porter developed a framework for understanding how individual firms can evaluate their costs and value-add to products and services, and then extend that framework to include assessments of complementary assets. The influence of networks and "embeddedness" from sociology (Granovetter, 1984) criticized the neoclassical view of the abstract economic actor in favor of an interpretation that accounts for the social relations present in all economic activity. The ideas of embeddedness were taken up in economic geography with Dicken and Thrift (1992), who encouraged this focus on the network relations of firms.

Perhaps the most direct contribution to the GPN framework is the concept of a "global commodity chain," developed by Gereffi and Korzeniewicz (1990, 1994) within economic sociology. As defined by Gerrefi et. al (1994: 2), global commodity chains are: "sets of interorganizational networks clustered around one commodity or product, linking households, enterprises, and states to one another within the world-economy. These networks are situationally specific, socially constructed, and locally integrated, underscoring the embeddedness of economic organization."

For Gereffi, the global commodity chain (GCC) concept was an effort to build on the "world-systems" categorization of economies with a more global focus that could inform economic

development policy, especially for developing countries. However, the GCC framework used the relatively gross spatial scales employed by the world systems work, of core, semiperiphery, and periphery. Another perceived shortcoming of GCC is an overdue emphasis on the state and its boundaries, leading to a kind of "methodological nationalism" (Hess and Yeung, 2006). Partly in response, Gereffi and others later evolved the global commodity chains into "global value chains," which drew more explicitly on business transaction cost economics to inform firm strategies around complementary assets (i.e., integration vs. outsourcing) and firm learning (Gereffi et al., 2005). The global value chain framework also described a typology of possible inter-firm relationships among members of the value chain. The relationships are to be determined by the complexity of information, degree of codification, and capabilities of the respective firms, and are labeled, in order of increasing dependency of the supplier on the primary firm: Markets, Modular, Relational, Captive, and Hierarchy (Gereffi et al., 2005).

The global production network (GPN) framework tries to incorporate influences from all of these antecedents in order to more accurately understand the increasingly complex and multi-dimensional networks of economic actors within which firms, states, non-governmental actors, and consumers operate. The GPN framework is characterized by five areas of focus: networks, value, power, embeddedness, and non-firm institutions (Hess and Coe, 2006; Henderson, et. al, 2002), which makes it an appropriate framework analyzing platform-based industries such as the smartphone industry. The emphasis on standards and non-firm actors, for example, is especially germane to platform ecosystems, given the fundamental role standards play at the technology layer. While a complete GPN analysis of the smartphone industry is outside the scope of this work, we do leverage key concepts from the GPN construct to guide our spatial argument around the app economy, especially power relations and value capture.

The framing of inter-firm power relations is useful in understanding the ability of the platform owner, i.e. Apple or Google, to establish governance policies and processes that are relatively inclusive or exclusive for external firms. As platform owners, they must design policies that incentivize app developers to innovate on their platform, while still maintaining control over the ecosystem; this "generativity paradox" can be seen as an ongoing, reflexive process (Tilson et. al,

2012). The power dynamic is not always as clear cut, however, as developers are sometimes able to push for more favorable terms, and the act of governance can be fluid (Manner et. al, 2013; Ghazawneh and Henfridsson, 2013).

A significant part of the power relationships, of course, is the ability to capture value from the ecosystem. We use value capture as a lens in multiple ways: At the firm level, we analyze the platform firms' public filings and announcements to determine the broad categories of revenue generation that they depend on for sustaining growth (e.g., the almost complete reliance by Google on advertising) in order to correlate our assumptions about user lock-in and value capture (Kenney and Pon, 2011). In a more dynamic fashion, we use the concept of value capture to guide our analysis of "digital flows" of apps and money through the app stores and across national borders (Pon et. al, forthcoming); by restricting the analysis to those apps that are likely to be earning revenue, we avoid sampling that large body of app development that is done as a hobby or experimentation, and therefore not relevant for our study (VisionMobile, 2014).

GPN analysis therefore has some degree of shared conceptual framing with the platform literature, and indeed can be seen as a parent or umbrella framework of which a technology platform is a sub-type. While it doesn't consider many of the strategic issues specific to platforms, its general constructs are relevant and valuable for the study of platforms, especially insofar as they help to develop a spatial dimension to a platform analysis.

The platform as market: conceptualizing the app stores

A significant, and growing, part of the smartphone industry is the app economy—the development and sale of apps and digital content through the app stores, virtual marketplaces managed by the platform owners. The two main stores, Apple App Store and Google Play, are the core mechanism by which most developers distribute and monetize their apps, and most end-users discover, download, and pay for apps (Heitkoetter, 2012). When Apple launched the iPhone in 2007, 3rd-party developers were not permitted to create native apps. Perhaps bowing to pressure from developers, Apple opened up the platform by releasing an SDK and launching the App Store in July

2008; there were 552 apps available at launch.[5] In early 2015, the App Store has 1.4 million apps, and has paid out a cumulative $25 billion to developers ($10 billion of that in 2014).[6] The rise of Google Play is no less spectacular. It too launched in 2008, with only a dozen apps,[7] but now hosts 1.4 million, growing 50% year-on-year in 2014 and now featuring over 400,000 developers.[8] While these two stores are by far the most dominant globally, there are also stores for each of the other major platforms—Amazon Appstore, Windows Phone Store, BlackBerry World, and Firefox Marketplace—as well as hundreds of 3rd-party stores that typically feature either Android apps or an assortment of apps from multiple platforms.

These platform-based app stores are examples of a "two-sided market," a conceptualization that has grown from the economics literature on network industries, including the early work by Katz and Shapiro (1985), David and Greenstein (1990), and Farrell and Saloner (1985; 1986), whose work on standards and network effects attempted to model participation and innovation outcomes given network effects and increasing returns. More recently, economists have focused explicitly on the multi-sided or two-sided market model, trying to determine, for example, the optimal pricing for different sides of the market, and for platform incumbents vs. new entrants (Rochet and Tirole, 2003; Parker and Van Alstyne, 2005; Weyl, 2010). This view assumes that the value for participating in the platform for either of the user groups is positively correlated with the amount of participation by the other side; that is, the more end-users of a platform, the more attractive it is for developers or suppliers to join, and vice versa (Boudreau, 2007; Evans and Schmalensee, 2006).

Of course, reaching a critical mass and successful equilibrium presents a classic chicken-or-egg challenge: the platform is only useful if both (or all) sides are at the table, yet convincing both sides to participate simultaneously without any evidence or guarantee of participation by the other side can be challenging (Evans, 2009). One way that platform firms address this challenge is through providing

[5] Lex Friedman, "The App Store turns five: A look back and forward" *Macworld*, July 8, 2013. http://www.macworld.com/article/2043841/the-app-store-turns-five-a-look-back-and-forward.html
[6] Apple press release January 8, 2015, "App Store Rings in 2015 with New Records"; http://www.apple.com/pr/library/2015/01/08App-Store-Rings-in-2015-with-New-Records.html
[7] Melissa Perenson, "Google Launches Android Market" TechHive, October 22, 2008. http://www.techhive.com/article/152613/google_android_ships.html
[8] AppFigures data: http://blog.appfigures.com/app-stores-growth-accelerates-in-2014/

incentives to one or more sides; for example, establishing different pricing schemes, or subsidies, for different user groups (Rochet and Tirole, 2003; Economides and Katsamaka, 2006). For example, in the video console market, end-users are subsidized (they can purchase a console for close to actual cost) while developers are charged licensing and royalty fees in order to have access to the platform (and thus, end-users) (Johns, 2006). However, these economic examinations of multi-sided markets are based on abstract models, mostly focused on pricing, and ignore the broader social and technological context in which these platforms-as-markets function. To understand the functioning of the app stores, we have to take a spatial perspective that examines how these proprietary markets shape, and are shaped by, the location of those who participate, and how this changes as the goods traded move from the physical to the virtual.

The market in (digital) space

From the Greek agora to the Moroccan bazaar, the market has historically been associated with a physical location as the site of exchange activity (Knorr Cetina, 2006; Geertz, 1978). This form of "central market" was dominant for hundreds if not thousands of years, and can still be seen in our local farmers' markets and flea markets (Sayer, 1995). As regional and global trade increased in the 16th century, new and more sophisticated supply chains and intermediaries gave rise to trade-based markets, which leveraged travel and transport costs to create value for consumers and producers (Knorr Cetina, 2006). Then more recently, industrialization and increased global trade led to the creation of the mass consumer market, embodied by the retail store: a dedicated interface between not only producers and consumers, but a host of intermediaries, suppliers, and now advertisers (Knorr Cetina, 2006).

But this spatial embeddedness of markets, as richly described by Geertz (1978)—where, for example, location matters not only at the scale of the market itself, but also for vendor location within the market, due to information asymmetries—has been transformed by advancements in ICT. This has happened in two ways: first, the digitization of existing markets, most notably the stock exchanges, has supplanted physical trading and exchange rooted in space with digitized communications from

actors that can be located virtually anywhere, what Knorr Cetina (2006) calls "scopic systems": "electronic and informational mechanisms of observing and contextualizing market reality and of back-projecting this reality onto the computer screens of globally operating traders and financial units." Of course, stock exchanges are unlike other types of markets discussed here in that they do not comprise the actual exchange of goods that can be consumed, but instead financial instruments. And despite stock markets' detachment from a central place for physical exchange, participation in stock markets is decidedly spatially organized, as firms make choices about which exchanges (i.e., which countries) they want to be listed on, and investors make choices about which firms (i.e., located where) they want to purchase stock in; this latter phenomenon manifests in investors being more likely to invest in firms that are located geographically closer to themselves (Wojcik, 2009). And while ICT continues to compress distance and time into ever smaller and more inscrutable units for these global scopic systems, even the millisecond and meter matter: high-frequency stock traders have reaped millions in profits through strategically siting their data servers closer to the stock exchange than rivals' servers, because even a few meters of physical distance translates into additional milliseconds of time to execute a trade, and those traders with the closest and fastest servers could outperform others sited further away.[9]

The second way ICT has transformed place-based markets is through the digitization of goods themselves, that is, the production and consumption of information goods in digital form (Quah, 2003; Moriset and Malecki, 2009). Markets based on these digital goods, including Apple's iTunes and the Android platform, exhibit not only Knorr Cetina's scopic characteristics, but also a fundamentally distinct system of production, distribution, and consumption based on computerization, digital formats, and electronic transmission. For example, while digital goods[10] may have high initial costs of

[9] As documented in the book *Flash Boys.* Even though all trades occur electronically, these traders are able to use this tiniest of differences to conduct arbitrage with incoming trades and actual market prices. Even in what may be the most virtualized and abstracted markets in existence—online stock exchanges are doubly abstracted in that they are not only non-physical, they trade securities instead of actual goods—place and distance matter. See http://www.telegraph.co.uk/finance/newsbysector/banksandfinance/10736960/High-frequency-trading-when-milliseconds-mean-millions.html

[10] We use "digital good" as a more precise alternative to "information good," as the latter can technically include product mediums such as pen and paper, or vinyl records. It is when this information is digitized into binary computer code of 1's and 0's that the full extent of advancements in ICT drive radical change.

production, marginal or reproduction costs are minimal and even close to zero, [11] and transportation costs across digital networks are similarly minimal[12] (Shapiro and Varian, 1999; Evans et. al, 2006). On the one hand, the seeming omnipresence of digital communication devices and networks—there are more mobile connections than people—seems to increase accessibility to these digital information goods from anywhere, anytime, leading Moriset and Malecki (2009) to liken these goods to Weber's "ubiquities," in the sense of universally accessible inputs. Yet as those authors rightly point out, access to communication networks and digital goods is not universal, and instead reflects the existing spatial socioeconomic realities such as state-provided infrastructure (Moriset and Malecki, 2009). And neither is the production of these information goods evenly distributed—the main developers of software and internet content more generally are most likely to be located in the United States and other Western regions with longer histories of ICT development (Zook, 2001; 2005).

The rapid growth of online commerce and new internet-facilitated business models around the turn of the century led to a surge in studies exploring the dynamics and potential impact of these emerging "digital," "electronic," or "online" markets. Focused primarily on e-commerce, these studies compared online and offline businesses in terms of pricing (Smith et. al, 1999), transaction efficiencies (Bakos, 1998), search and discovery (Brynjolfsson et. al, 2003), and the reconfiguration of the value chain (Sarkar et. al, 1998). However, these studies primarily examine how internet technologies are changing business models and processes for existing products—for example, how the sale of CDs or books online, instead of in a physical retail store, creates new value chains, lowers prices for consumers, and enables greater options in the market (Brynjolfsson et. al, 2003).

Therefore the emergence of the Android platform, and other network-based markets for digital information goods, represents a break of the physical siting of the market from a central place. Yet these markets are still socially embedded in their production systems and consumption patterns, which

[11] The fact that digital goods have minimal reproduction and distribution costs precludes standard pricing theory based on costs of production; firms instead have to create new metrics for valuing and pricing goods (Shapiro and Varian, 1999).

[12] Of course, the actual costs incurred for transmission can vary widely both in absolute as well as relative terms for the end-user—for many digital goods, the seller assumes the buyer has unlimited bandwidth, while for low-income users in infrastructure-poor areas, data transmission costs can be prohibitive and directly impact the viability of digital goods sales.

are strongly shaped by the global geography of ICT development. Essential market functions—the facilitation of economic exchange between buyers and sellers, and the supporting institutions that enable this exchange, are mediated by this new market form and its scopic systems, yet function in similar ways. The information-seeking behavior described by Geertz (1978) is accomplished through an online search; the face-to-face relationship-building as a precursor to trust as described by Moriset and Malecki (2009) has been substituted for online reviews and "trusted developer" icons; and the mass market retail store described by Knorr Cetina (2006) is now virtualized in an online marketplace, complete with advertisements and other intermediaries.

Part II. “Structuring the Smartphone Industry: Is the Mobile Internet OS Platform the Key?”

This section was previously published as an article in the Journal of Industry, Competition and Trade. Microsoft Word version copied here.

Kenney, Martin, and Bryan Pon. "Structuring the smartphone industry: is the mobile internet OS platform the key?" Journal of Industry, Competition and Trade 11.3 (2011): 239-261.

Structuring the Smartphone Industry: Is the Mobile Internet OS Platform the Key?*

Received: January 27, 2011

Martin Kenney
Department of Human and Community Development
University of California, Davis
mfkenney@ucdavis.edu
&
Senior Project Director
Berkeley Roundtable on the International Economy

and

Bryan Pon
Ph.D. Student
Geography Graduate Group
University of California, Davis

* The authors acknowledge the assistance, discussions, and comments of Andrew Hargadon, Greg Linden, Jonathan Murray, Petri Rouvinen, Timo Seppälä, Urs von Burg, and John Zysman. We are particularly thankful to Joel West who contributed so much to the first version of this paper and without whose penetrating ideas and deep knowledge of the industry, we could never have completed the project. This research is a part of the ongoing collaboration of BRIE, the Berkeley Roundtable on the International Economy at the University of California at Berkeley, and ETLA, The Research Institute of the Finnish Economy.

Keywords: Platforms, Industry structure, Smart phones, Android, iPhone

Abstract

With the launch of the original iPhone, Apple redefined the "smartphone" product category and accelerated the convergence of traditional mobile telephony, Internet services, and personal computing into a new industry. As these sectors merge into a single device, formerly separate industry architectures and their constituent firms are being forced into direct competition. We test theories of industry architecture and technological platforms regarding their ability to explain the strategies of key entrants in navigating the transition. We analyze in detail the actions and strategies of four major competitors, including Apple, Google, Microsoft, Nokia, and, more briefly, Research in Motion and HP/Palm, from the framework of technological platform theory. Our analysis suggests that currently some competitors are following traditional platform strategies, but that Google and Apple appear to have adopted strategies at odds with platform literature. We examine how the dynamics of this convergence may lead to a reconsideration of certain tenets of platform theory.

Introduction

The convergence of mobile telephony, Internet services, and personal computing is resulting in the emergence of the smartphone and the "mobile Internet" (Ishii 2004; Funk 2001). Information and communications technology (ICT) firms that were formerly operating in only one of these sectors have been drawn into a new competitive landscape that is collapsing software, hardware, and services in new ways. The key firms competing in this new market have each been a market leader in their original industry, and therefore bring previously successful core competencies to their smartphone efforts. Analysis of the extent to which firms are trying to leverage their existing assets and strategies to capture market share and value in the smartphone industry may be instructive in understanding how ICT firms create and control value in new industries. For many of the firms entering the smartphone market, capturing a portion of the total value created by the industry is believed to be a key to future growth and profits.

The interest is understandable. Today more than 1.3 billion mobile phone handsets are being sold annually, and in 2010 smartphones made up almost 20% of that total (Gartner, 2010; Ahonen,

2010). In contrast to standard mobile phones, "smartphones" are powerful computing devices offering traditional wireless voice service as well as native software applications and, perhaps most importantly, the ability to connect to and run a myriad of Internet-based services including email, geo-location, streaming video, and social networking, while providing a good user experience. Sales of smartphones are increasing almost 100% per year, and total global sales volume is expected to surpass that of PCs by 2012 (Gartner, 2010). By collapsing the boundaries between previously distinct devices, smartphones are subsuming sales of mobile phones entirely and, increasingly, netbook and notebook PCs. To complicate the landscape, the smartphone is not the only device at stake, tablets and ebook readers are emerging as key components of the mobile universe. Across all devices, total mobile revenues—including advertising, subscriptions, handsets, applications, and so on—are forecast to surpass $1 trillion by 2014 (Gartner, 2010). Given the rate at which smartphone are penetrating the market and component prices are declining by 2015 there will be, at least, 2 billion smart mobile devices in use globally.

The nature of the smartphone device and industry lends itself to analysis from a technology platform perspective. Scholars and management consultants have identified platform control as a key feature for business success in the ICT industries. Michael Cusumano, drawing upon his studies of Microsoft, Cisco, and Intel, concluded that the winner of technological competitions is "often who has the best platform strategy and the best ecosystem to back it up (Cusumano 2010: 34)." The opportunity to establish platforms often comes in the early phases of an industry's development or when a major technological/market discontinuity occurs. De novo firms and previously existing firms can use the discontinuities to enter a space and displace or subordinate previous incumbents or, even more cleverly, use new business models to transform the value capture equation. This paper explores the competition in the emerging smartphone operating system race, looking at value capture and customer lock-in strategies through the lens of industry architecture and platform theory. Other potentially important market characteristics were considered but not explicitly included. For example, the potential for anti-trust legislation or other legal challenges to firms' bundling of software, hardware, services, and content could alter strategies. Given the anti-trust lawsuits against Microsoft for bundling its browser to its operating system, it seems reasonable to expect regulatory scrutiny