Loyola University Chicago Law Journal

Volume 45
Issue 2 *2013 Winter*

Article 7

2013

Composition of Innovative Activity in ICT Equipment R&D

Yasin Ozcan
Kellogg School of Management, Northwestern University

Shane Greenstein
Kellogg School of Management, Northwestern University

Composition of Innovative Activity in ICT Equipment R&D

*Yasin Ozcan and Shane Greenstein**

Has the market structure for inventive ideas in the Information and Communications Technology ("ICT") equipment industry undergone dramatic changes in the last three decades in the United States? What does statistical evidence from U.S. patent activity suggest about change to the concentration of sources of inventive ideas? This Study characterizes levels, and changes in those levels, in the concentration of sources of new invention from 1976 to 2010. The analysis finds pervasive deconcentration across a wide set of areas. It also finds that the deconcentration takes place despite the role lateral entry by existing firms plays in driving concentration levels up. Furthermore, the evidence suggests that the deconcentration trend cannot be attributed to a single supply factor in the market for ideas, such as the breakdown of AT&T during the deregulation of the telecommunications industry.

TABLE OF CONTENTS

INTRODUCTION 480
I. ICT EQUIPMENT INDUSTRY CONCENTRATION 486
  A. *Historical Overview* 487
  B. *Theoretical Framework* 490
II. DATA 492
  A. *Patent Sample Selection* 493
  B. *Concentration and Other Measures* 495
III. DECONCENTRATION OF PATENT OWNERSHIP 498

* Ph.D. Candidate and Professor, respectively, Kellogg School of Management, Northwestern University. y-ozcan@kellogg.northwestern.edu, greenstein@kellogg.northwestern.edu. We thank Colleen Mallahan for research assistance and Ben Jones and seminar participants and the editors for comments. We also thank the Dean's Office and the GM Strategy Research Center of the Kellogg School of Management for funding. We are responsible for all errors.

A. Composition of Ownership in New Patents: Historical Trends 498
B. Composition of Ownership in New Patents: The Model.........500
CONCLUSION 502
TABLES, FIGURES, AND APPENDICES...504

INTRODUCTION

It is not an exaggeration to say that the market structure for the Information and Communications Technology ("ICT") equipment industry in the United States has undergone enormous changes in the last three decades. This Study focuses on one aspect of these changes: namely, the concentration of origin of innovative ideas—a new method, idea, or product. Thirty years ago, most innovation took place in established firms—particularly large, centrally controlled laboratories, such as Bell Labs and IBM Labs. While such activity continues, analysts noticed decades ago that such labs had lost their prominence to widespread, decentralized, and small-scale innovators.[1] This observation about the origins of inventive activity goes by many names in many models.[2] This Study uses the label "divided technical leadership" ("DTL").[3] DTL plays a key role in models of open innovation[4] and in models of open and proprietary platforms.[5] It also plays a key role in models of the externalization of research and development ("R&D") by large firms that use acquisitions of smaller firms for many of these innovative activities.[6] Firms such as Cisco, IBM, and Apple participate

1. Richard S. Rosenbloom & William J. Spencer, *Introduction: Technology's Vanishing Wellspring*, *in* ENGINES OF INNOVATION: U.S. INDUSTRIAL RESEARCH AT THE END OF AN ERA 1, 3 (Richard S. Rosenbloom & William J. Spencer eds., 1996) ("Research activities have been downsized, redirected, and restructured in recent years within most of the firms that were among the largest sponsors of industrial research.").

2. Rosenbloom and Spencer discuss the importance of research from laboratories, but also the rise of small scale research in start-ups and other young firms. *See id.* Sometimes this is called "Silicon Valley Style" R&D, or "entrepreneurial" R&D. Related labels are discussed in the text.

3. Timothy F. Bresnahan & Shane Greenstein, *Technological Competition and the Structure of the Computer Industry*, 47. J. INDUS. ECON. 1, 1–40 (1999) (discussing the market structure concept of DTL).

4. HENRY CHESBROUGH, OPEN INNOVATION: THE NEW IMPERATIVE FOR CREATING AND PROFITING FROM TECHNOLOGY 43 (2003).

5. Shane Greenstein, *Innovative Conduct in U.S. Commercial Computing and Internet Markets*, *in* HANDBOOK ON THE ECONOMICS OF INNOVATION 477, 493 (Bronwyn Hall & Nathan Rosenberg eds., 2010).

6. Joshua S. Gans et al., *When Does Start-Up Innovation Spur the Gale of Creative Destruction?*, 33 RAND J. ECON. 571, 571 (2002).

in such activities, each having made more than 100 acquisitions over the last two decades.

While the trend has received notice, examination of its causes has not moved much beyond casual empirism and anecdote. This Study addresses the absence of statistical information and econometric analysis by providing an examination of the long run causes behind DTL. It examines whether statistical evidence of long-term changes shows a deconcentration of sources of inventive ideas, as held by conventional models of DTL.[7] This Study also provides the first characterization of levels and changes in those levels in the concentration of origins of innovation in the ICT equipment industry. To construct measures of the origins of innovation, this Study examines the concentration in granted patents in ICT equipment from 1976 to 2010.[8] The data reveal very large changes over time, which motivates a second question: what are the determinants of deconcentration? This part of the Study uses variance between different segments to identify determinants of changes in concentration. The statistical exercise measures the contribution of economies of scope, product market leadership and entry by domestic and foreign firms from 1976 to 2010.

This Study utilizes a data set constructed from XML and text files of patents granted by the U.S. Patent and Trademark Office ("USPTO") between 1976 and 2010. The newly constructed data covers four more recent years than National Bureau of Economic Research ("NBER") patent data files, the standard data source for many studies on patents.[9] There are several reasons to use this data. ICT equipment is an important downstream market, involving hundreds of billions of dollars of investment by end users. The creation of equipment upstream is a knowledge-intensive and patent-intensive activity, accounting for roughly 14% of all U.S. patents.[10] This data also suits the goal of

7. By deconcentration we mean a change in the composition of the sources of inventive ideas, away from a small number of firms to many more. We will propose a way to implement how to measure deconcentration in this Study.

8. The "granted patents" of this Study will encompass all those issued by the U.S. Patent and Trademark Office ("USPTO"), and the date of issuance will be the date associated with their grant.

9. For details on the NBER patent data files, see Bronwyn H. Hall et al., *The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools* 3 (Nat'l Bureau of Econ. Research, Working Paper No. 8498, 2001) [hereinafter *Data File*] ("The goal of this paper is to describe the data base on U.S. patents."). This is a widely used source of patent data, primarily because prior researchers have cleaned the files of misspellings and other errors, and made them compatible with other common sources of data.

10. This figure is the ratio of the number of patents in Information and Communications Technology Equipment ("ICTE") Industry to the number of patents in all technology classes in the U.S., covering the period from 1976 to 2010. The patents for the ICTE Industry consist of patents

analyzing long-term trends and factors more comprehensively than prior research; as explained in the text, the length of time covered is novel.

The findings first cover long-term trends. The Study reveals a grain of truth to conventional belief about the presence of DTL. We show that a deconcentration trend is present in the ownership of new patent applications. Interestingly, the analysis documents considerable variation in the size and scope of the changes; while some segments of ICT equipment have undergone dramatic changes in concentration, others have undergone less dramatic change. Overall, however, the data reveal a dramatic decline in concentration. While on average the top twenty-five firms accounted for 72% of the new patents in 1976, the top twenty-five firms accounted for only 55% of the new patent grants by the end of the sample. Furthermore, this trend of deconcentration is even starker when the sample is restricted to high-quality patents based on the citations received by each patent, with a decline in top twenty-five firms' ownership from 86% to 62% over the same period.[11]

Why does this deconcentration arise? On the supply side, large firms in some industries may be utilizing economies of scope by entering new technical areas for their invention, which may appear as increased or decreased concentration depending on the size of the entry.[12] This Study uses such lateral entry as a proxy for economies of scope, and, as shown below, finds evidence that ownership concentration increases with lateral entry. In addition, this Study provides evidence that de novo firm entry, which may be used as a proxy for reduced transaction costs of entry by inexperienced firms, accounts for part of this deconcentration, but not all types of entry. Importantly, this Study rejects the notion that non-U.S. firm entry caused the change, which is important to test because the U.S. economy began to involve a larger fraction of imports and exports over this period.[13] Rather, established changes in concentration may come

---

in forty-four (44) technology classes based on the USPTO classification system. More information on the data is provided in the Data Section and in the Appendices.

11. As we explain in the text, a high-quality patent is one that receives a number of "forward" citations, namely, citations from patents that come later, and the number of forward citations places the patent among the top quartile of all patents receiving citations.

12. In the context of production, "economies of scope" is usually defined as lowering the average cost of producing two or more products. In the context of invention, "economies of scope" refers to the lowering of the average costs of inventing in two or more technical areas.

13. As with the rest of the literature, we are somewhat cautious in our interpretation of foreign firms. A patent owned by Sony, for example, will appear as a U.S. patent due to the location of its U.S.-based subsidiary. As with the prior literature, *see, e.g.*, Bronwyn H. Hall, *Exploring the Patent Explosion*, 30 J. TECH. TRANSFER 35 (2005) [hereinafter *Patent Explosion*], we focus on changes due to U.S. patents with U.S assignees and non-U.S assignees, and examine whether the surge in patenting with non-U.S assignees accounts for change.

from two distinct areas of the ownership distribution: (1) declines in the leading, large firms, and (2) an increase of innovation in small and entrepreneurial firms within the U.S. These entry results are consistent with the growth of small and entrepreneurial firms as a source of ideas.

Evidence for the decline in the importance of large firms is more mixed. This Study provides evidence that decreases in product market leadership explain the deconcentration in some instances. However, the preponderance of evidence suggests this is not the single most important factor explaining variance across technical areas, which we define below. More specifically, this Study reveals that long-term trends in deconcentration cannot be fully accounted for by the divestiture of AT&T,[14] or the loss in commercial leadership at IBM, Motorola, or any other large firm in the industry.[15] Hence, this Study rejects the most sweeping version of the hypothesis that points to one antitrust case, one company's strategic error, or the break-up of one large, leading innovator of yesteryear as the cause for this change in structure.

This Study relates to the research streams in two main channels. First, the deconcentration of ownership relates to the literature on DTL, as noted, and more broadly, debates about the causes of market leadership in innovative activities. Following this literature,[16] this Study generally distinguishes between product market leadership and technological leadership, and focuses on the latter. Second, this is the first Study to investigate the scope and cause of deconcentration in innovation in the ICT equipment industry specifically. Finally, this Study differs from prior literature with its focus on understanding the causes behind changes in technical leadership in the more recent decades, an unexamined question in prior work. This Study establishes this novelty with a more detailed comparison and contrasts with prior research.

This Study also builds on prior research into patenting. The dramatic increase in U.S. patenting activity since the 1970s has attracted the

---

14. As discussed below, the divestiture of AT&T broke AT&T into seven regional local telephone companies and one other firm that combined a long distance telephone company with a telephone equipment firm. It also divided Bell Labs into two entities, one for the long distance/equipment firm, and another for the consortium of local firms. After this division Bell Labs began to decline in size.

15. As discussed below, IBM and Motorola both had leading positions as technical pioneers and providers of computing and cellular telephone equipment, respectively. Over the course of the period covered by our patents both firms would lose market share in output markets.

16. This literature is reviewed in Greenstein, *supra* note 5. Technological leadership refers to activities—principally innovation—that lead a firm to introduce and implement a new product or service. Commercial leadership refers to activities—chiefly production and distribution—that lead a firm to gain larger market share at the expense of rivals.

attention of scholars: Kortum and Lerner investigated the U.S. patenting activity, finding that changes in U.S. patent policy—specifically the establishment of the U.S. Court of Appeals to hear patent cases—did not have a verifiable impact on the increased patenting activity.[17] Instead, Kortum and Lerner associated the increase in patenting to an increase in U.S. innovation and changes in the management of R&D, which may have included actions such as reallocating efforts to more applied problems with higher patent yields.[18] Kim and Marschke analyzed the same issue, concluding that the increased patenting activity resulted from increases in R&D in some sectors, and increases in the patent yield in the computing, electronics, and auto sectors.[19] Hall found that growth occurred in complex product industries such as telecommunications, where products are based not only on a single patent, but on multiple and technically complementary patents.[20] She further concluded that increased patenting activity by firms in complex product industries spilled over to those firms' patenting behavior in other industries, resulting in an overall increase in patenting activity across all technology classes.[21]

These studies mostly focused on activity at the technology class level as described in NBER patent data files.[22] In contrast, this Study focuses on the distribution of the increase in patenting between firms within each technology class, or in other words, differs by looking at the technical leadership of firms in addition to the main trends in the technology class level. Furthermore, the up-to-date data on patents allow this Study to answer the open questions suggested by earlier studies, including: "What happened during the 1990s? Did the positive premium for entry with patents continue during the rapid growth of the computing and electronics sector in the late 1990s? Has the growth in patenting continued to be due almost entirely to U.S. firms in computing and

17. Samuel Kortum & Josh Lerner, *What Is Behind the Recent Surge in Patenting?*, 28 RES. POL'Y 1 (1998).

18. *See id.* at 4 ("[T]he increase in patent activity here seems to be a consequence of a worldwide increase, along with a recent improvement in the relative performance of U.S. inventors.").

19. *See* Jinyoung Kim & Gerald G. Marschke, *Accounting for the Recent Surge in U.S. Patenting: Changes in R&D Expenditures, Patent Yields, and the High Tech Sector*, 13 ECON. INNOVATION & NEW TECH. 543 (2004).

20. Bronwyn H. Hall et al., *Market Value and Patent Citations* 28 (Nat'l Bureau of Econ. Research, Working Paper No. 7741, 2000), *available at* http://escholarship.org/uc/item/0cs6v2w7.

21. *See Patent Explosion*, *supra* note 13, at 46.

22. For details on the NBER patent data files, see *Data File*, *supra* note 9.

electronics?"[23]

This Study supports the view that the changes in ownership concentration are consistent with a trend towards divided technical leadership, namely, more widespread access to the fundamental knowledge and building blocks for innovative activity in this sector of the economy. This is the framework put forward in Bresnahan and Greenstein, which suggests that platforms in an industry with rapid technological development can be quite concentrated if leading firms retain proprietary rights over standards.[24] In the face of technological competition between platforms, however, it is possible for different firms to take the leading position on different segments of the platform.[25] Hence, high rates of firm entry and exit may occur without changing the concentration at the platform level.[26] This Study contributes to the literature by providing considerable evidence consistent with the central premise: that DTL has increased over time in inventive ideas upstream to computing and related sectors, such as Internet equipment, both prior to the period analyzed by Bresnahan and Greenstein and continuing thereafter.

Furthermore, the results of deconcentration can be interpreted as a switch from a so-called Chandlerian to a so-called Schumpeterian market in terms of market leadership in innovative activities in ICT equipment.[27] The Chandlerian view of market leadership focuses on the accumulative nature of leadership, and asserts that market leadership persists for a long time, embedded within the organizational form of leading firms. The Schumpeterian "creative destruction" view states that market leadership

---

23. *See Patent Explosion*, *supra* note 13, at 15 (noting that electrical and computing technology firms in the U.S. accounted for the higher rate in overall patenting between 1980 and 1989).

24. Bresnahan & Greenstein, *supra* note 3, at 1.

25. A platform in computing is a reconfigurable base of compatible components on which users build applications. Interoperable standards typically play the role in making components compatible and can be part of a larger strategy to govern how different groups interact with one another. *See* Greenstein, *supra* note 5, at 497.

26. For example, there was considerable entry into the provision of components in the personal computer ("PC") platform commonly referred to as "IBM-compatible" PCs, and turnover in leadership in many of its key components, such as word processing, spread sheet, and presentation software. Yet, the competition "between" platforms—mainframe, minicomputer and PC—hardly moved at all in response to changes "within" platforms. *See* Bresnahan & Greenstein, *supra* note 3, at 1 (emphasizing "the importance of technological competition between computer 'platforms,' not firms").

27. Franco Malerba & Luigi Orsenigo, *Schumpeterian Patterns of Innovation Are Technology-Specific*, 25 RES. POL'Y 451, 452 (1996). For a detailed discussion of Schumpeterian and Chandlerian views of market leadership, and a benchmark for the *transient* and *long time* periods, see John Sutton, *Market Share Dynamics and the 'Persistence of Leadership' Debate*, 97 AM. ECON. REV. 222 (2007).

is transient and subject to frequent threats of replacement. Malerba and Orsenigo address this debate by identifying two classes of sectors. The first class of sectors, the Schumpeter Mark II (Chandlerian) sectors, has high concentration of innovative activities, big innovator size, high rank stability among innovators, and low entry levels. These sectors include chemicals and electronics firms. The second class of sectors, the Schumpeter Mark I (Schumpeterian) sectors, has the opposite characteristics, and includes mechanical technologies and traditional sectors. This Study shows that the ownership of innovative activities was highly concentrated in the early periods of the data, which is consistent with Malerba and Orsenigo's classification of electronics as a Schumpeter Mark II industry.[28] In addition, this Study further shows that this high concentration of ownership has seen a dramatic change over the last four decades, and, using Malerba's and Orsenigo's framework, may be interpreted as a switch from a Chandlerian to a Schumpeterian market structure.

The rest of this Study is organized as follows: Part I provides background on the history of the ICT equipment industry, which motivates a framework for analysis. Part II describes the data construction, sample selection, and variable construction. Part III presents the empirical methodology and deconcentration of patent ownership results in ICT equipment. The Study concludes that the evidence points towards a substantial reduction in concentration. The Appendices contain details of data construction and data linking methodology.

I. ICT Equipment Industry Concentration

How should deconcentration of innovation activity in the ICT equipment industry from the late 1970s to the present be characterized? ICT equipment plays an important role in markets for electronics, computing, and infrastructure of radio, television, voice, and broadband communication services. It would take several books to describe the changes in market structure during this time, and this Part cannot hope to review all the details.[29] The purpose here is only to refresh the reader's memories about what the literature takes for granted about major changes

28. Malerba & Orsenigo, *supra* note 27, at 454.

29. The scale of change is too large to summarize in a short article. A comprehensive review would involve a history of every leading firm (e.g., IBM, AT&T, and Motorola), a review of every major change at the federal level in regulatory and antitrust policy, a review of an extraordinary range of technical developments linked to the emergence of cellular telephony, packet-switching (e.g., the Internet), and the personal computer, as well as miniaturization of digital technologies (e.g., semi-conductors) and its consequences for a range of equipment markets.

in the concentration of origins of inventive ideas in a wide set of related industries. This will provide just enough of a brief overview to guide the development of a framework for the statistical exercise.

A. Historical Overview

Prior to the 1980s, the ICT equipment industry consisted of various segments, depending on whether it was oriented towards computing, as it was then understood, or towards communications, namely, voice or data.[30] Both of these segments were highly concentrated in final goods markets.[31] At the end of the 1970s, IBM dominated the computing segment with its mainframe systems and components built around those systems.[32] It also dominated the personal computer system market for a short time, growing a small systems division that in 1984 would have been the third largest computer company in the world (behind Digital Equipment Corporation and IBM itself).[33]

Starting in the mid-1980s and accelerating thereafter, IBM lost market share in personal computers and in many of the peripheral markets.[34] After the introduction of the IBM personal computer ("PC") in 1981, a wide range of firms entered into printers, software, component production, and local area networks.[35] In the 1990s, Microsoft and Intel began to assert control over an increasing fraction of valuable components within the PC market; nonetheless, a large number of firms played a role in many of its segments.[36]

30. NAT'L RESEARCH COUNCIL, INNOVATION IN INFORMATION TECHNOLOGY 19–20 (2003), *available at* http://www.nap.edu/catalog/10795.html.

31. JAMES W. CORTADA, INFORMATION TECHNOLOGY AS BUSINESS HISTORY: ISSUES IN THE HISTORY AND MANAGEMENT OF COMPUTERS 80 (1996).

32. *Id.* at 82, 169.

33. *Id.* at 177–78.

34. *See* WILLIAM APRAY & MARTIN CAMPBELL-KELLY, COMPUTER: A HISTORY OF THE INFORMATION MACHINE 139 (1996) ("IBM enjoyed an exceptional dominance of computing for at least two decades. But such dominance could not last forever, and it didn't."); Bresnahan & Greenstein, *supra* note 3, at 26–27; Greenstein, *supra* note 5, at 514 ("At the outset of the 1990s, before the Internet commercialized, IBM's mainframe business had begun to decline significantly. This led the board to remove the CEO and break with precedent by hiring a CEO from outside the company . . . .").

35. *See* Greenstein, *supra* note 5, at 499.

36. For more on platforms in computing markets and the ecosystems that grew up around them, see, e.g., ANNABELLE GAWER & MICHAEL A. CUSUMANO, PLATFORM LEADERSHIP: HOW INTEL, MICROSOFT AND CISCO DRIVE INNOVATION 15–16 (2002); ANDREW S. GROVE, ONLY THE PARANOID SURVIVE: HOW TO EXPLOIT THE CRISIS POINTS THAT CHALLENGE EVERY COMPANY AND CAREER 105–06 (1996). *See generally* DAVID G. MESSERSCHMITT & CLEMENS SZYPERSKI, SOFTWARE ECOSYSTEM: UNDERSTANDING AN INDISPENSIBLE TECHNOLOGY AND INDUSTRY (2003) (describing the software industry as a complex ecosystem made up of numerous,

Before the 1980s, AT&T was the dominant provider of networking equipment in the voice segment, largely due to its regulated monopoly position in telecommunication services; approximately 90% of AT&T's equipment purchases were supplied from its equipment subsidiary, Western Electric. The voice segment was based on circuit-switching technology and provided the infrastructure for local and long-distance telephone companies.[37] Furthermore, AT&T fought regulations that ended its requirement that any equipment attached to its network had to be supplied by AT&T, even on the end-user site.[38] The purchase behavior and network attachment requirement of AT&T restricted entry into the telecommunications equipment markets, thus carrying AT&T's dominant position in telecom services into the telecom equipment sector.[39]

These fights yielded change, but slowly. In 1968, AT&T lost an antitrust suit against Carterfone Company, and was forced to permit private interconnection equipment on the AT&T network.[40] In 1975, the Federal Communications Commission ("FCC") extended the *Carterfone* decision to all private subscriber equipment that was registered to and certified by the FCC.[41] These decisions enabled entry into the telecommunications equipment industry; as long as AT&T remained the dominant purchaser of equipment, however, entry was limited.[42] The market structure changed further with the 1974 U.S. Department of Justice antitrust suit against AT&T.[43] The case was settled in 1982, with AT&T divesting its local telephone service into seven independent, regional holding companies, breaking up equipment purchasing decision-making. As a result, the telephone markets underwent considerable

complementary systems); U. VON BURG, THE TRIUMPH OF ETHERNET: TECHNOLOGICAL COMMUNITIES AND THE BATTLE FOR THE LAN STANDARD (2001).

37. An introduction on this topic is provided in Ana Aizcorbe et al., *The Role of Semiconductor Inputs in IT Hardware Price Decline: Computers vs. Communications*, *in* HARD-TO-MEASURE GOODS AND SERVICES: ESSAYS IN HONOR OF ZVI GRILICHES. 351, 351–79 (Ernst R. Berndt & Charles R. Hulten eds., 2007), *available at* http://www.nber.org/chapters/c0883.pdf.

38. For a discussion of these issues, see JONATHAN E. NUECHTERLEIN & PHILIP J. WEISER, DIGITAL CROSSROADS: AMERICAN TELECOMMUNICATIONS POLICY IN THE INTERNET AGE 57–59 (2005); PETER TEMIN, THE FALL OF THE BELL SYSTEM: A STUDY IN PRICES AND POLITICS 85, 97, 99–112 (1987).

39. TEMIN, *supra* note 38, at 58.

40. *In re* The Use of the Carterfone Device in Message Toll Telephone Service, 13 F.C.C.2d 420, 422 (1968).

41. United States v. Am. Tel. & Tel. Co., 62 F.C.C.2d 1102, 1119–20 (1975).

42. *Id.*

43. *See* NUECHTERLEIN & WEISER, *supra* note 38, at 60–64; TEMIN, *supra* note 38, at 104–12.

changes in the early to mid-1990s.[44]

The data segment was based on packet-switching technology and supplied the communication equipment required in the computing industry, including modems and local area networks. Until the emergence of the Ethernet standard, this segment was characterized by proprietary protocols.[45] Only with widespread use of the Ethernet standard in the late 1980s and the Internet IP stack in the early 1990s did non-proprietary standards begin to shape industry structure.[46]

The networking and Internet revolution of the 1990s blurred the distinction between different segments of ICT equipment. This process sometimes receives the label "convergence," which means that previously independent product market segments increasingly become substitutes or complements in demand.[47] On the computing side, systems of PCs and workstations were initially connected with a local area network ("LAN"). Over time, client-server systems within large enterprises and across ownership boundaries were established. Novell, 3Com, Oracle, and Cisco were among the firms with dominant positions in this era.[48] With widespread Internet use, the scope of ambitions became quite large, touching on virtually every economic activity in which transmission of information played an important role.[49] This period was marked by economic experiments across a wide range of activities that overlapped with applications of computing and communications, as well as any related upstream or downstream activity. It was marked by optimism and labeled "the dot-com bubble" in recognition of the many startups that ended with the top-level domain name "com."[50]

In contrast, by the beginning of the millennium, many layers of the industry had undergone upheaval.[51] Some of this was associated with

44. *See* ROBERT W. CRANDALL & LEONARD WAVERMAN, TALK IS CHEAP: THE PROMISE OF REGULATORY REFORM IN NORTH AMERICAN TELECOMMUNICATIONS 5–7, 9–10 (1995) (discussing the results of the phone monopoly break-up and the fragmentation that has resulted).

45. *See* VON BURG, *supra* note 36, at 3 (contrasting the development of Ethernet, an open technological standard, to proprietary technology, over which the developer exerts exclusive control).

46. *See* Greenstein, *supra* note 5, at 505–29.

47. *Id.*

48. *See* Bresnahan & Greenstein, *supra* note 3, at 29 (noting Novell's and Oracle's strength in their respective markets).

49. *See* DALE W. JORGENSON ET AL., PRODUCTIVITY, INFORMATION TECHNOLOGY AND THE AMERICAN GROWTH RESURGENCE (2003).

50. *See generally* Greenstein, *supra* note 5.

51. *See* Brent D. Goldfarb et al., *Searching for Ghosts: Business Survival, Unmeasured Entrepreneurial Activity and Private Equity Investment in the Dot-Com Era* 2–3 (Robert H. Smith

large, painful adjustments due to a decline in demand that was linked to the implementation of the 1996 Telecommunications Act[52] and the resulting growth and telecom meltdown. Some of it was due to the bursting of the dot-com bubble.[53] Eventually, the equipment market stabilized, leaving Cisco in the dominant position in enterprise computing to serve data communications. Yet other firms that grew spectacularly during the 1990s, such as JDS Uniphase, Corning, Lucent, Nortel, and 3Com, did not fare as well.

This brief review suggests several of the core questions that motivate the statistical work of this Study. First, is the evidence consistent with the common presumption that there has been a deconcentration in the ownership of innovative ideas? Second, can this deconcentration be explained by something straightforward, such as the divestiture of AT&T, the loss of commercial leadership at IBM or Motorola, or any other large industry firm? Third, what role do other factors play, such as firm entry, particularly non-U.S. firm entry, which has accelerated over this period? And fourth, has the externalization of R&D by established firms merely changed the structure of the origins of innovation, but not its concentration as it relates to final output markets?

B. Theoretical Framework

This Section provides a brief overview of the framework of this Study. It fixes, introduces, and establishes a few key ideas, and provides a roadmap for later developments.

Following prior literature,[54] this Study divides the industry into an upstream sector that supplies invention and a downstream sector that supplies products. The inventions can take a variety of forms, such as implemented ideas in prototype products or patent filings or copyrighted designs. The downstream sector employs inventions from the upstream sector in production.

The literature on the rise of DTL focuses on the increasing infrequency of situations where one firm has a monopoly over an idea.[55] In practice, these ideas come from very specific classes of technologies and map into very specific product markets, such as specific component markets or

---

Sch. of Bus., Paper No. RHS-06-027, 2005), *available at* http://ssrn.com/abstract=825687 (providing an extensive discussion of examples).

52. Telecommunications Act of 1996, Pub. L. No. 104-104, 110 Stat 56 (codified as amended in scattered sections of 47 U.S.C.).

53. Goldfarb et al., *supra* note 51, at 6–7.

54. ASHISH ARORA ET AL., MARKETS FOR TECHNOLOGY: THE ECONOMICS OF INNOVATION AND CORPORATE STRATEGY 6 (2001); Gans et al., *supra* note 6, at 583.

55. *See, e.g.*, Bresnahan & Greenstein, *supra* note 3, at 3–5 (discussing DTL).

software applications. The literature stresses that such monopolies are less likely to arise where many technical substitutes can emerge—for example, as when multiple firms can supply a microprocessor to perform basic memory functions. Substitutes are more likely to emerge in settings where many potential inventors generate similar ideas, and where entry into production of ideas is less costly. The latter situation characterized many consumer equipment markets, such as cameras, televisions, and video players. It also characterized many software markets for personal computers in the 1980s and 1990s, and, later, many Internet application markets in the late 1990s and beyond.

There are many alternative ways of measuring settings where many potential inventors generate similar ideas. For reasons explained below, this Study settled on a top-twenty-five concentration ratio over the ownership of inventive ideas, which is labeled as C25, in technological class, which is indexed by i. Illustrating the concept, a technological class i is said to be more concentrated if the largest twenty-five firms own 80% of the inventive ideas instead of, say, 50% of the ideas in that technology class.

The literature discussed many related measures of concentration for a sector, and these are book-ended by two concepts: one related to the *flow of new ideas*, and another related to the *stock of ideas*. The existing literature on DTL suggests the flow of ideas is relevant for fostering entry into product markets, for example, while the stock of ideas is relevant for new combinations of technologies fostering entry or industrial change. This Study focuses on the flow of ideas concept, as this is the first step towards implementing these concepts. The latter is left for later work.

The first key question concerns changes in concentration of ownership over time: Is the evidence consistent with decreasing concentration over time? When looking at a concentration ratio, for example, that will focus on the question for each technical sector, labeled i, namely:

$$(C25_{flow})_{it} - \left(C25_{flow}\right)_{it-1} < 0\,.$$

Generally, as discussed below, a wide range of technology classes did become more deconcentrated. That motivated the second question, concerning the causes of changes in concentration over time. In general, this approach will identify causes of the variance in changes of concentration between different technology classes. That is, this Study posits:

$$(C25_{flow})_{it} - \left(C25_{flow}\right)_{it-1} = f(Supply\ in\ i, demand\ in\ i).$$

The literature on DTL frames the open question: what factors caused changes in concentration? As the review of the history discussed above suggests,[56] important supply-side factors include the decline of dominant firms, increasing economies of scope across technology sectors, the entry of foreign firms, and the entry of small firms. Important demand-side factors include the increasing use of mergers by leading firms to obtain invention from external sources, increasing acceptance of technical products from unbranded firms by users, and the increasing use of open standards that permit customers to buy interoperable products from more than one supplier. This Study will construct measures for all three supply factors, while the demand factors will be absorbed into time trends, for reasons described below.

II. DATA

Patents are one of the most utilized sources of information in the innovation literature.[57] The use of patent data as a proxy for economic activity dates back to Schmookler and Griliches, and since then, an extensive literature on using patents as indicators of innovative activity has developed.[58] This Study keeps with this literature and focuses on patents granted in the ICT equipment industry as a proxy for the origins of innovative activity.[59] Since pursuing questions related to DTL led us to modify the practices underlying existing, widely-used patent datasets, we first explain our overlap with and departures from the existing literature.[60] We then establish changes in the level of ownership composition of new innovative activity, and then link these changes to

---

56. *See supra* Part I.A.

57. *See, e.g.*, Zvi Griliches, *Patent Statistics as Economic Indicators: A Survey*, 28 J. ECON. LITERATURE 1661, 1661–62 (1990) (discussing the benefits of studying patents, including their availability and objectivity); Sadao Nagaoka et al., *Patent Statistics as an Innovation Indicator*, *in* HANDBOOK OF THE ECONOMICS OF INNOVATION, *supra* note 5, at 1085, 1085–86 (discussing the basic characteristics of patent data as an innovation indicator).

58. *See* Griliches, *supra* note 57, at 1662–63; Nagaoka et al., *supra* note 57, at 1085–86 ("Recently, patent information is increasingly used to analyze innovation and the innovation process, and patent statistics are increasingly used as a measure of innovation."); Jacob Schmookler, Invention and Economic Development (1951) (unpublished Ph.D. dissertation, University of Pennsylvania) (on file with author).

59. Using patenting to measure invention has one principal advantage: it provides a standardized measure over a very long time period. There is, as yet, no other feasible way to measure the extent and direction of inventive activity over four decades across a set of related technical segments, as done in this Study.

60. *See infra* Part II.A.

underlying supply-side factors, which include new entry, lateral entry (a firm's economies of scope), and growth.[61]

The standard source for patent data in the innovation literature has been the NBER patent data file.[62] In contrast, we used raw USPTO files to construct an updated patent data file. Appendix I describes the construction of patent data from 1976 to 2010.

A. Patent Sample Selection

The ICT equipment industry is a knowledge-intensive market that corresponds to hundreds of billions of dollars in investments by end users in the downstream, and roughly 14% of U.S. patent stock in the upstream. We identify the ICT equipment industry in the patent data by extracting forty-four patent technology classes from the newly constructed USPTO patent data: fourteen technology classes identified as communications by Hall, Jaffe, and Trajtenberg;[63] twenty-two technology classes in the 700 ranges; and eight classes identified as relevant to telecommunications in the USPTO communications report. We then drop fourteen classes due to sparse patenting activity.[64] The classification variable is taken from the December 2010 version of the U.S. Patent Grant Master Classification File ("MCF")[65] published by the USPTO.[66]

While our patent data include granted patents between 1976 and 2010, we encounter truncation created by the application-grant lag in the patent system.[67] Accordingly, we restrict our sample to patents applied for between 1976 and 2007. The final dataset has 550,884 patents with

---

61. *See infra* Part II.B.

62. *See* Griliches, *supra* note 57, at 1662 (discussing studies); Nagaoka et al., *supra* note 57, at 1112–17 (same).

63. *Data File*, *supra* note 9, at 41.

64. Appendix III contains lists of all considered classes. The focus of this Study is the ICTE Industry, and these forty-four classes are identified as parts of ICTE in widely accepted USPTO classifications and NBER patent data classifications. The dropped classes correspond to roughly 10% of the entire patenting in ICTE.

65. The latest version of the MFC can be found at https://explore.data.gov/Business-Enterprise/Master-Classification-File-MCF-Patent-Grant-Patent/vg9q-x87u.

66. The USPTO organizes patents into approximately 450 technology classes, and 150,000 subclasses, based on common subject matter, in which a class delineates one technology from another. For more information on the USPTO classification system, see OFFICE OF PATENT CLASSIFICATION, http://www.uspto.gov/patents/resources/classification/index.jsp (last visited Oct. 31, 2013).

67. Some patents applications in 2005 and 2006, for example, had not yet been granted as of 2010, and, therefore, we could not examine them. The phenomenon became more severe as we approached 2010.

primary technology classes in the thirty ICT equipment classes, assigned to 38,359 unique assignees.

The 550,000 patents granted in the ICT equipment industry during our sample period correspond to roughly 14% of all patenting activity in the United States. Figure 1 provides a breakdown of granted patents over the years. As observed, the number of patents granted in ICT equipment follows a trend akin to the total number of utility patents granted by the USPTO: the number of patents granted increases starting in the 1980s, followed by a sharp decline in the 2000s due to the patent grant delay—the time between the patent application by inventors and their receipt of a grant from the USPTO. The figure also provides the relative magnitude of unassigned patents, roughly 30,000, which we drop from our sample, as we are interested in analyzing the assigned patents.[68] Given the small magnitude, it is unlikely that the unassigned patents drive any of our results.[69]

The patent literature firmly establishes that patent values are highly skewed, with studies noting that the most valuable 10% of patents account for as much as 80% of total value of patents.[70] Below we provide results for all patents, and then for patents that receive the bulk of citations, which are presumed to be of higher quality,[71] and, relatedly, the

---

68. The dropped, unassigned patents are held by independent inventors. An independent inventor is defined by the USPTO as "a person whose patent, at the time of grant, has ownership that is unassigned or assigned to an individual (i.e., ownership of the patent is not assigned to an organization)." USPTO Patent Tech. Monitoring Team, *Independent Inventors By State By Year All Patent Types Report*, U.S. PAT. AND TRADEMARK OFF., http://www.uspto.gov/web/offices/ac/ido/oeip/taf/inv_all.htm (last modified April 2, 2013).

69. Unassigned patents belong to independent inventors, and calculating ownership concentration in independent inventors involves identifying unique inventors in the patent data in a similar method to our firm name-linking algorithm described in the Appendices. Therefore, the unassigned patents are dropped from our sample. It is unlikely that any independent inventor would have enough patents to be among the top twenty-five firms, and therefore this drop impacts our concentration measure described in Part II.B *infra* only at the denominator. Given the small size of the unassigned patents, and their disbursement across technology classes, including them in the data would simply pull down our concentration measure by a negligible amount.

70. *See, e.g.*, F. M. Scherer & Dietmar Harhoff, *Technology Policy for a World of Skew-Distributed Outcomes*, 29 RES. POL'Y 559, 559 (2000). For other studies stressing the skewed distribution of patent values, see Dietmar Harhoff et al., *Citation Frequency and the Value of Patented Inventions*, 81 REV. ECON. & STAT. 511 (1999); Ariel Pakes & Mark Schankerman, *The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources*, *in* R&D, PATENTS, AND PRODUCTIVITY 73 (Zvi Griliches ed., 1984).

71. An interpretation of this approach is through the Schumpeterian framework. *See* JOSEF SCHUMPETER, THE THEORY OF ECONOMIC DEVELOPMENT (Redvers Opie trans., 1934) (distinguishing inventions and innovations: an invention is a potential innovation, and becomes an innovation only when it is commercialized). One could argue that the count of all patents is a better proxy for inventions and the count of high-quality patents is a better proxy for innovations. *See*

most valuable.[72] We define high-quality patents as the top quartile within their technology class-year group cells in terms of citations received.[73]

B. Concentration and Other Measures

This Section describes the market structure and technology supply proxies we use in our empirical framework. Table 3 provides a summary of these variables.

Our main variable is the patent ownership concentration in a technology class. We capture the ownership concentration of granted patents in each technology class-year group as the share of top firms in the ICT equipment industry. More specifically we create variables $C1_{flow}, C2_{flow}, \ldots, C25_{flow}$, where CX_{flow} is the share of patents by the top X firms within the technology class-year group. In each year group we reselect the top firms; in other words, even though the number of firms used to calculate CX is kept constant at X, the set of firms may be different from period to period. We stop at $C25_{flow}$ because in many of the technology class-year groups, the top twenty-five firms reach 100% ownership in the early years of our sample. Table 3 reports that on average the top twenty-five firms in a technology class-year group own 60% of patents. Though, as discussed in the next Section, there is considerable variation in this concentration over time.[74]

Firm entry into innovative activities provides one theory on deconcentration.[75] In an effort to capture the impact of firm entry, we have three classes of entry variables. In the first class, patent-weighted entry level is constructed by two measures of entry based on the previous patenting activity of the firm, which we label as *new entry* and *lateral entry*, and which we define below. Firm *i* is considered a new entrant to technology class *j* in period *t* if the firm does not have any patents in any of the ICT equipment classes prior to period *t*, and has at least one patent

Joel West & Marcel Bogers, Profiting from External Innovation: A Review of Research on Open Innovation 26 (September 13, 2011) (unpublished manuscript), *available at* http://ssrn.com/abstract=1949520 (discussing patents as a measure of innovation); *see also* Ricardo J. Caballero & Adam B.A. Jaffe, *How High Are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth* 7 (Nat'l Bureau of Econ. Research, Working Paper No. 4370, 1993) (offering another alternative interpretation "that patents are proportional to ideas, and that citations are proportional to ideas used").

72. *See, e.g.*, Harhoff et al., *supra* note 70, at 511 (discussing the value of citations); Hall et al., *supra* note 20, at 24, 31–34 (finding a significant relation between the value of patents and the number of citations they receive).

73. We have also examined the top decile without any large change in inference.

74. *See infra* Part III.A.

75. Kortum & Lerner, *supra* note 17, at 1–22.

in technology class *j* in period *t*. When such an entry occurs, we consider all patents of firm *i* in period *t* in technology class *j* to be patents by a new entrant, and calculate the new entry share by dividing the total number of new entry patents by the total number of patents in technology class *j* in period *t*. The new entry variable then captures the level of transaction costs of entry into the supply of ideas, particularly from those who previously had made none, where the transactions costs of entry would matter most for outcomes. When we restrict the variable to account for only foreign entry, we then capture the transaction costs of entry by non-U.S. firms. Table 3 suggests that firms that had no prior ICT equipment innovation activity produce, on average, 12% of patents in a technology class. This share increases to 19% when the sample is restricted to high-quality patents.[76]

In addition to firms entering into the ICT equipment industry from outside, firms may also be active in one ICT equipment class and later move to a new ICT equipment class. We consider such firms as lateral entrants. More specifically, we consider firm *i* a lateral entrant to technology class *j* in period *t* if the firm did not have any patents in technology class *j* prior to period *t*, had at least one patent in another ICT equipment technology class prior to period *t*, and had at least one patent in class *j* in period *t*. We then calculate the *lateral entry share* as the ratio of patents by lateral entrants in period *t* in class *j* to the total patent count in period *t* in class *j*. We theorize that a higher lateral entry level implies higher economies of scope across different technology classes. The summary statistics in Table 3 reveal that on average, 11% of patents come from lateral entrants, with the share going up to 15% for high-quality patents.

The two entry variables, *new entry share* and *lateral entry share*, proxy for patent-count weighted entry into a technology. We should note that when combined, these two variables capture the inverse of the serial dependence of patenting by firms already in a technology class. In other words, considering the 11% new entry and 12% lateral entry averages, we deduce that on average, 77% (100–11–12) of patents come from firms that already had patents in a technology class in prior periods. As a result, when we include both entry variables in the model, we also account for the serial dependence.[77]

76. When we restrict the sample to high-quality patents, the entry variables capture entry into the high-quality patent pool rather than entry into the entire patent pool. In other words, a firm with many low-quality patents and no high-quality patent in prior periods would be considered an entrant in the first period it produces a high-quality patent.

77. Serial dependence (or serial correlation) is the dependence of the value of a variable in a

The second class of entry variables is the growth in the number of firms active in a technology class. Using simple firm counts, we calculate the growth in the number of firms over time. We see that on average the number of firms has increased by 13% every two years, with firms located outside the U.S. having a relatively higher growth rate of 17%.[78] As an overwhelming majority of the firms in the sample are U.S.-based, the total growth in the number of firms is very close to the growth in U.S.-firms, which is around 13%.

The growth in the number of issued patents constitutes our third class of independent variables. We see that on average the patent count has grown by 19% every two years (18% in domestic and 26% in foreign firms). When we take into account the 13% average increase in the number of firms over two-year periods, which is considerably less than the 19% growth in patent count, we deduce that patent growth is coming from both entrants and incumbents.

The final class of control variables in our model consists of proxies for increase or decrease in product market leadership: dummies for the presence of a big firm. As the history of this sector suggests, the breakdown of AT&T might have been responsible for the deconcentration in patent ownership, and that conjecture calls for these statistical controls. In an attempt to discern whether the existence of big firms, namely AT&T, Motorola, and IBM, has an impact on the concentration, we include lagged indicators for their existence among the top five patent applicants. We see that the presence of AT&T is somewhat dwarfed by the strong presence of IBM: IBM is among the top five patent applicants in 41% of technology class-year group cells, whereas AT&T and Motorola are in the top five patent applicants in only 37% and 25% of the cells, respectively.[79]

time on the value of the same variable at another time. For more information, see generally WILLIAM H. GREENE, ECONOMETRIC ANALYSIS (4th ed. 2000).

78. One should take the statements about foreign firms with a grain of salt for the following reason. The *foreign* indicator in the patent data captures the location of a firm, but not its origin or ultimate ownership. For example, even though practitioners would consider Sony Electronics Inc. a non-U.S. firm, in the patent data it is located at Park Ridge, New Jersey, and therefore is considered to be a U.S. firm (e.g., USPTO patent 5,828,956).

79. These cells correspond to two-year periods as opposed to one year. This construct is explained *infra* in Part III.A.

III. Deconcentration of Patent Ownership

A. Composition of Ownership in New Patents: Historical Trends

This Section describes long-term trends, which characterizes our endogenous variable. We construct a measure of concentration, and then analyze the new patent creation across the thirty ICT equipment technology classes. To capture the dynamics of new patent creation, we calculate the *patent flow* variable—the number of new patents a firm has applied for in a given year and was granted at a later date.[80]

Patent applications do not get approved with any regularity, however, and sheer randomness can lead to little activity in some technical classes for extended periods of time. To ensure that we have enough observation in each bucket (i.e., each patent class each period) in our analyses of patent flow, we use two-year intervals as the measure of time instead of individual years. Therefore, the observation level throughout the patent flow analyses is a technology class-year group.

We use $C25_{flow}$, the share of top twenty-five firms in new patents, as our measure of concentration. In calculating the $C25_{flow}$ measure, we reselect the top firms in each period.[81] We choose $C25_{flow}$ as opposed to other CX_{flow} values because in many cells, $C25_{flow}$ reaches to 100% for the early periods of our sample. We discuss the choice of the concentration measure further in Part III.B.[82]

Figure 2 illustrates the CX_{flow} values for technology class 385 (Optical Waveguides). The top line in Figure 2 represents the share of top twenty-five firms in the class ($C25_{flow}$), and the bottom line represents the share of the top firm only ($C1_{flow}$). The share of the top twenty-five firms has seen a decline from around 70% in 1976–77 to around 41% in 2006–07. In fact, we observe a similar trend in twenty-six of the thirty classes in our sample. In only four classes the values of $C25_{flow}$ fluctuate. All these trends suggest a deconcentration of ownership in new patents in our sample period.

80. The patent grants may come many years after a patent is applied for, and this delay is coined as the patent application-grant delay. The convention in the literature on patents is to use the patent application year as the year of the innovation/invention because the application year is closer to the actual creation of the idea; whereas the delay, hence the grant year, is a function of other factors including the workload and staffing issues at the USPTO. In this Study, we follow this convention, and use the patents applied for and granted between 1976 and 2010.

81. To be clear, the measure is based on twenty-five firms in each period, even if there are changes in the identities of these firms from period to period.

82. *See infra* Part III.B.

We now turn to Table 1 to observe this deconcentration trend across all technology classes. Table 1–A shows the distribution of $C25_{flow}$ values across all technology classes for all assigned patents in the ICT equipment sample. The mean value of the top twenty-five firms' new patent share across technology classes follows a gradual decline over the years from 72% in the 1976–77 period to 55% in 2006–07. When the sample is restricted to the high-quality patents—the top 25% of patents within each class-year group based on citations received—we observe an even sharper reduction in concentration from 86% in 1976–77 to 62% in 2006–07.[83]

Figure 3 is simply an alternative way of observing this trend of deconcentration: in 1976–77, five classes possessed more than 90% of the new patents, whereas in 2006–07 no classes showed such concentrated ownership at the top. The difference is even starker for high-quality patents: when we restrict the sample to the top 25% of patents, the top twenty-five firms in 1976–77 possessed more than 90% of the new patents in ten classes, as opposed to only one class in 2006–07.[84]

We now do our first investigation of the potential causes of this deconcentration across technology classes. Did the divestiture of AT&T in 1982 play a role? To see if this claim holds in a first pass through the data, we calculate a simple statistic, the number of firms that contribute 90% or more of the changes in $C25_{flow}$, the share of top twenty-five firms, over our sample period. The results are presented in Table 2: Panel A reports the changes for all ICT equipment patents, and Panel B reports the same analyses for high-quality patents. We see that of the twenty-six classes with deconcentration, in only two classes are three or fewer firms responsible for 90% or more of the reduction in $C25_{flow}$. In the remaining twenty-four classes, there is an industry-wide deconcentration trend, which suggests that the breakdown of AT&T, or another leading firm, cannot be the sole reason for the established deconcentration. The qualitative observations remain the same when we restrict the sample to high-quality patents.

83. Similar results hold when we restrict the sample further to include only the highest-quality patents, those in the top 10%: the average value drops from 96% in 1976–77 to 71% in 2006–07.

84. Similar results hold for the top 10% of the patents, with twenty-four classes in 1976–77 and six classes in 2006–07.

B. Composition of Ownership in New Patents: The Model

Part III.A presented historical trends and provided evidence for a deconcentration trend in the ownership of new patents in the ICT equipment industry.[85] This Section combines these historical trends into a single fixed effects model to provide a coherent framework of the potential causes of the established deconcentration. In this analysis, $C25_{flow}$, the share of top twenty-five firms is the dependent variable. The basic model is as follows:

$$\begin{aligned}(C25_{flow})_{jt} = {} & \beta_1 * (\text{New Entry})_{jt} + \beta_2 * (\text{Lateral Entry})_{jt} \\ & + \beta_3 * (\text{Growth})_{jt} + \beta_4 * \delta_{j,t-1,AT\&T} \\ & + \beta_5 * \delta_{j,t-1,,Motorola} + \beta_6 * \delta_{j,t-1,IBM} + \gamma_j + \theta_t + \varepsilon_{jt},\end{aligned}$$

Here, *j* is the technology class indicator and *t* is the time indicator. The list of regressors include new entry and lateral entry into technology classes, growth measures, and indicator variables for the presence of big firms, namely AT&T, Motorola, and IBM.[86] We use two sets of growth measures: one for growth in the number of firms, and a second for growth in the number of patents. We further divide these growth variables into two components: growth in U.S.-based firms and patents, and their foreign counterparts. The growth measures are highly correlated (the Pearson correlation between total firm growth and total patent growth is 0.83).[87] Therefore, we use either the firm-based or the patent-based measure in a single model.

We present the results of the fixed effects models in Table 4, Panel A. The dependent variable in the model is $C25_{flow}$, the share of the top twenty-five firms in new patents. All models include class fixed effects; models 1–4 include a linear and a quadratic time trend, whereas models 5–8 include time fixed effects. The standard errors are clustered by technology-class. The columns differ in the inclusion of different patent growth and number of firm growth variables.

85. *See supra* Part III.A.

86. The construction and summary statistics of these variables are provided *supra* in Part II.B and *infra* at Table 3.

87. The Pearson correlation is a measure of linear correlation between two variables, and gives values between -1 and 1. A Pearson correlation of zero indicates that the two variables are not linearly dependent, and as the correlation moves away from zero in either direction, the linear dependence between the variables increases. Furthermore, a positive Pearson correlation value indicates that as one variable increases so does the other, and a negative correlation implies that as the value of one variable increases the other decreases.

The main qualitative results seem to hold across all models, and here we provide illustrations using the results from column 1. Growth in the number of firms is one of the main drivers of deconcentration. The Table shows that a 1% growth in the number of firms results in a decrease of 5.8% in the ownership share of top twenty-five firms. This is large. A technology class at the average firm growth rate of 13% every two years faces a reduction of approximately 0.8% (which equals 5.87% * 0.13) in the share of top twenty-five firms in two years, even after controlling for individual class effects and time trend.[88] When we break the growth variable into U.S.-based growth and foreign growth, we observe that contrary to conventional wisdom, only the U.S.-based growth is a driving force of deconcentration, and the foreign growth does not have a statistically or economically significant impact on our concentration measures.

Neither new entry share nor growth in the number of patents, however, seems to have a statistically significant impact at the 90% confidence level, though the sign of the estimates are in the negative direction, as expected. The lateral entry is associated with an increase in the ownership of top firms, and the impact is both statistically and economically significant: a technology class experiencing the average level lateral entry, 11% per period, faces a 3.5% (which equals 30.37% * 0.11) increase in $C25_{flow}$. This result may be driven by the fact that firms conducting lateral entry operate in multiple segments of the industry, and hence are expected to have a bigger operation than others.[89]

Finally, the models suggest that the existence of AT&T as one of the top five patent owners in the prior period does not have a statistically significant impact on the concentration of the patent class, which is consistent with the earlier trend analyses. The coefficient of the IBM indicator is also not significant. The presence of Motorola as a prior top-five patent applicant, however, is associated with an approximately 1.5% increase in the ownership concentration of the patent class over two years. A detailed look at Motorola's activity reveals that it focuses on five technology classes in which the deconcentration is less than the average across all technology classes. It is unclear whether the increased concentration is driven by the presence of Motorola or whether it is

88. This result is consistent with the fertile technology hypotheses of Kortum and Lerner. *See* Kortum & Lerner, *supra* note 17.

89. An alternative explanation may be that the lateral entrants move from large technology classes to small technology classes, hence dominating the class they move into. A breakdown of patent counts by technology class indicates that no single technology class in our data sample dominates, hence we refute this alternative explanation.

simply an artifact of selection on technology classes in which Motorola used to provide equipment.

The econometric results across all models show that growth in the number of firms is an important driver of deconcentration, suggesting that a smaller transaction cost for entry results in lower ownership concentration. Lateral entry works in the opposite direction of entry by increasing the concentration of patent ownership. There are similar results for the growth of high-quality patents in the number of firms; the impact of lateral entry, however, is mitigated and also loses its statistical significance in some of the models (Panel B of Table 4).[90] Note that lateral entry in this context means having a high-quality patent in one ICT equipment class, and producing a new high-quality patent in another ICT equipment class in which the firm did not have high-quality patents previously; having low-quality patents in either industry has no effect on the entry measure among high-quality patents.

These findings also raise an interesting open question. Looking at how the new entry and lateral entry vary over time (averaged across technology classes), we observe a declining trend in both. The new entry share starts around 15% in 1978–79 and gradually drops to 6.4% in 2006–07. The lateral entry share follows a similar declining trend, with 21% in 1978–79, and 6.1% in 2006–07. It is possible that the factors of lateral entry and new entry only reflected a one-time change that has largely played itself out. If both have declined permanently, then neither factor can play as large a role in driving change going forward.

CONCLUSION

This Study characterizes long-term trends related to the concentration of the origins of inventive ideas in the ICT equipment industry. Analyzing the concentration in granted patents in this industry from 1976 to 2010, this Study compares measured changes against popular assumptions about the size and scale of changes in innovation.

Overall, this Study reveals a substantial decline in concentration. The data show that the deconcentration trend is present in the ownership of new patents and that the size and scope of the changes vary considerably, with some segments of ICT equipment undergoing much more dramatic changes in concentration.

---

90. In unreported results, these changes are even more pronounced when we restrict the patents to the top 10%: lateral entry is no longer statistically significant in any of the models, though the total growth in the number of firms is still of the same magnitude and is statistically significant. The results also hold qualitatively.

This Study also provides evidence about the causes of this change. The statistical evidence is consistent with explanations that stress the role of supply-side changes. This Study presents evidence that firm entry accounts for part of this deconcentration. Importantly, this Study rejects the notion that non-U.S.-firm entry caused the change, and the notion that one antitrust case, one company's strategic error, or the break-up of one large leading innovator accounts for this change in structure.

The results of this Study motivate open questions for future research. Firstly, this Study has examined changes in the ownership concentration of new patents. Patents, however, are valid for seventeen to twenty years, and many lose their relevance quickly. It is imperative to investigate how the changes in the flow of patents impact the long term changes in the stock of patents, and patents that are granted not only in the year of the flow, but also during the previous twenty years.

Furthermore, this Study uses proxies from the supply side of the upstream innovation markets, and does not explicitly model the demand for innovation. The demand for innovation, such as acquisition of startup firms by incumbents, or increasing product market demand, may also be playing a role in the observed long-term trends of deconcentration. For example, does the demand for acquisition of startup (i.e., the market for ideas) undo the deconcentration of ownership? This is possible by transferring the innovation stock from entrants to incumbents. It is one of the key open questions motivated by this first look at the long-term history.

Table 1: Distribution of $C25_{flow}$ Values
Panel A: All ICT Equipment Patents

Year Group	Mean (%)	St. Dev. (%)	10%	25%	50%	75%	90%
76–77	72	17	53	62	70	86	100
78–79	72	17	52	62	67	85	98
80–81	72	16	53	63	72	82	90
82–83	69	16	51	61	68	76	94
84–85	64	14	48	56	62	72	83
86–87	62	13	45	55	63	70	80
88–89	60	13	45	53	63	68	76
90–91	61	12	46	53	62	69	76
92–93	60	12	40	54	61	68	74
94–95	58	12	35	54	61	66	71
96–97	58	12	36	53	60	65	71
98–99	56	12	36	50	58	62	71
00–01	53	12	35	45	55	59	66
02–03	53	12	37	46	55	59	65
04–05	54	13	38	49	53	61	73
06–07	55	13	37	49	55	63	71

Notes: Evolution of the patent application *flow* share for top twenty-five firms that are ultimately granted on or before 2010. Each row corresponds to a two-year time period. The sample includes patent applications from thirty patent technology classes in the ICT equipment industry, at all levels of patent quality.

Table 1: Distribution of $C25_{flow}$ Values
Panel B: Top 25% of ICT Equipment Patents

Year Group	Mean (%)	St. Dev. (%)	10%	25%	50%	75%	90%
76–77	86	15	66	77	88	100	100
78–79	83	16	65	72	80	100	100
80–81	81	16	63	71	80	100	100
82–83	80	17	59	68	75	100	100
84–85	77	19	51	67	72	100	100
86–87	73	18	49	63	71	84	100
88–89	71	16	51	60	68	81	100
90–91	70	15	52	61	68	77	91
92–93	67	13	49	60	66	74	82
94–95	62	11	43	60	65	70	73
96–97	60	12	43	53	60	69	74
98–99	58	11	38	53	61	66	71
00–01	56	12	35	50	57	62	72
02–03	55	13	36	48	56	60	72
04–05	59	13	41	53	59	66	77
06–07	62	14	46%	53%	60%	73%	82%

Notes: Evolution of the patent application *flow* share for top twenty-five firms that are ultimately granted on or before 2010. Each row corresponds to a two-year time period. The sample includes patent applications from thirty patent technology classes in the ICT equipment industry, and the highest quartile of patents, where quality is measured by citations received.

Table 2: No. of Companies Accounting for 90% of Change in $C25_{flow}$

Panel A: All ICT Equipment Patents

No. of Companies	No. of Classes
1–3	2
4–19	14
20–24	10
Total	26

Notes: The number of ICT equipment industry patent technology classes that went through a deconcentration of patent *flow* ownership from 1976 to 2007, grouped by the number of companies that account for the 90% of the deconcentration. The sample includes all levels of patent quality.

Panel B: Top 25% of ICT Equipment Patents

No. of Companies	No. of Classes
1–3	3
4–19	13
20–24	11
Total	27

Notes: The number of ICT equipment industry patent technology classes that went through a deconcentration of patent *flow* ownership from 1976 to 2007, grouped by the number of companies that account for the 90% of the deconcentration. The sample includes the highest quartile of patents, where quality is measured by citations received.

Table 3: Summary Statistics of Key Patent *Flow* Variables

Variable	Mean (%)	Std. Dev. (%)
$C25_{flow}$	60	14
New Entry Share	12	8
Lateral Entry Share	11	7
Growth in No. of Firms		
Total	13	25
US only	13	27
Foreign only	17	54
Growth in No. of Patents		
Total	19	35
US only	18	37
Foreign only	26	81
Firm in Top 5 in Previous Period		
AT&T	0.37	0.48
Motorola	0.25	0.43
IBM	0.41	0.49

Notes: The sample includes patent applications from all levels of quality in the period 1976 to 2007 that are ultimately granted by the USPTO on or before 2010. The averages are across the thirty ICT equipment industry patent technology classes, and two-year time period cells. $C25_{flow}$ is the patent application share of top twenty-five companies within a cell. New Entry Share is the share of patents in a technology class in a period that are held by assignees that did not have any patents in any ICT equipment industry patent technology classes in prior periods. Lateral Entry Share is the share of patents in a technology class in a period that are held by assignees that had patents in other ICT equipment industry patent technology classes in prior periods, but did not have any patents in the current technology class in an earlier period. Growth is measured within each technology class across two consecutive two-year periods. The firm dummies indicate the presence of the firm among the top five patent *flow* holders in the previous two-year period.

Table 4: OLS Analysis of Patent *Flow* Ownership Concentration
Panel A: All ICT Equipment Patents

Dependent Variable: $C25_{flow}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Entry Share	-28.29 (34.95)	-29.66 (35.41)	-34.76 (34.88)	-34.95 (35.01)	-20.74 (35.74)	-22.59 (36.32)	-29.23 (35.90)	-29.58 (36.11)
Lateral Entry Share	30.37 (12.03) **	29.52 (12.52) **	25.41 (12.74) *	24.83 (13.13) *	35.71 (12.69) ***	34.9 (13.45) **	29.46 (13.51) **	29.22 (14.26) **
Total Growth in No. of Firms	-5.87 (1.95) ***				-7.87 (2.37) ***			
US only		-4.57 (1.57) ***				-5.98 (1.86) ***		
Foreign only		-0.16 (0.32)				-0.34 (0.36)		
Total Growth in No. of Patents			-0.69 (1.12)				-1.39 (1.44)	
US only				-0.28 (0.83)				-0.97 (1.04)
Foreign only				0.02 (0.2)				-0.03 (0.26)
Lagged Dummies if Firm Is in Top 5								
AT&T	-1.43 (0.95)	-1.26 (0.96)	-1.06 (0.95)	-1.01 (0.94)	-1.83 (0.97) *	-1.63 (0.97)	-1.31 (0.94)	-1.28 (0.94)
Motorola	1.55 (0.83) *	1.52 (0.84) *	1.52 (0.83) *	1.5 (0.83) *	0.74 (0.74)	0.75 (0.77)	0.9 (0.75)	0.9 (0.76)
IBM	-0.02 (1.15)	-0.16 (1.14)	-0.28 (1.06)	-0.34 (1.07)	-0.17 (1.31)	-0.31 (1.3)	-0.32 (1.2)	-0.37 (1.2)
Time Trend	-1.76 (0.60) ***	-1.82 (0.60) ***	-2.27 (0.66) ***	-2.33 (0.68) ***				
Time Trend Sq.	0.03 (0.03)	0.03 (0.03)	0.06 (0.03) *	0.06 (0.03) *				
Intercept	73.17 (6.82) ***	73.38 (6.84) ***	75.19 (6.56) ***	75.32 (6.64) ***	68.18 (7.90) ***	68.39 (8.02) ***	70.36 (7.68) ***	70.39 (7.78) ***
Time Fixed Effects	-	-	-	-	Yes	Yes	Yes	Yes
R-Squared	57%	56%	55%	55%	59%	58%	56%	56%

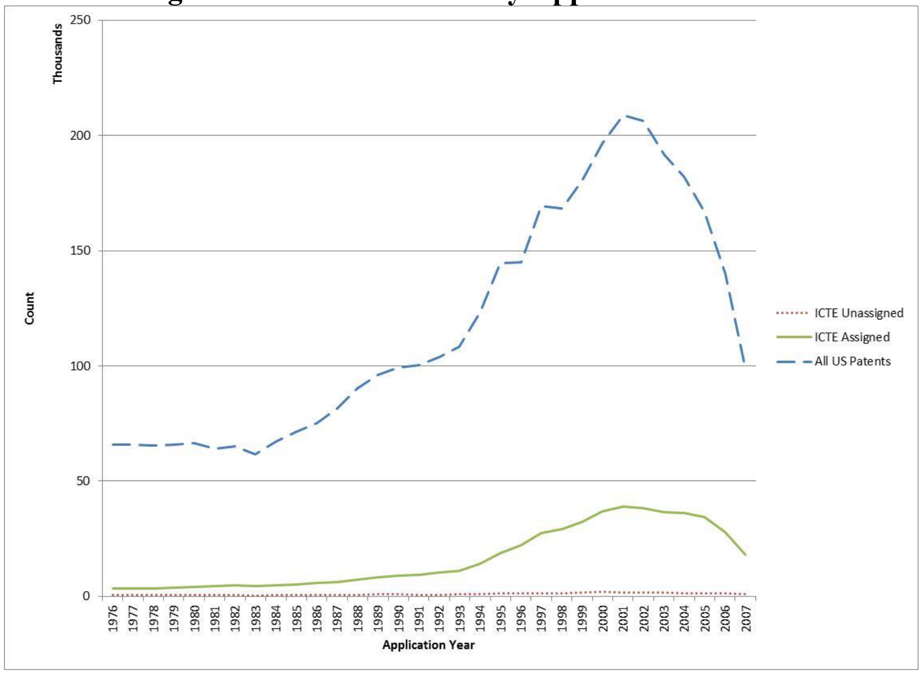
Notes: Regressions are ordinary least squares, with S.E. in parentheses. ***, **, * indicate significance at the 1%, 5%, and 10% levels respectively. Standard errors are clustered by class. An observation is a patent technology class and a two-year time period. N is 450. Each model includes technology class fixed effects. Models 1-4 include a linear and a quadratic time trend; models 5-8 include time fixed effects. The sample includes patent applications from all levels of quality in the period 1976 to 2007 that are ultimately granted by the USPTO on or before 2010.

Table 4: OLS Analysis of Patent *Flow* Ownership Concentration
Panel B: Top 25% of ICT Equipment Patents

Dependent Variable: $C25_{flow}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
New Entry Share	-1.45 (12.51)	0.89 (12.87)	-9.09 (13.4)	-7.59 (13.62)	-1.63 (11.72)	1.65 (12.2)	-8.62 (12.27)	-5.95 (12.71)
Lateral Entry Share	18.93 (9.77) *	20.28 (9.30) **	13.66 (9.74)	15.45 (9.3)	22.74 (9.95) **	24.94 (9.30) **	17.65 (9.77) *	20.32 (9.19) **
Total Growth in No. of Firms	-7.82 (1.76) ***				-6.43 (1.95) ***			
US only		-6.98 (1.15) ***				-5.96 (1.21) ***		
Foreign only		-2.29 (0.61) ***				-2.01 (0.66) ***		
Total Growth in No. of Patents			-3.22 (1.22) **				-1.34 (1.33)	
US only				-2.62 (1.19) **				-1.12 (1.27)
Foreign only				-1.57 (0.50) ***				-1.24 (0.52) **
Lagged Dummies if Firm Is in Top 5								
AT&T	0.27 (1.08)	0.16 (1.21)	0.32 (1.11)	0.16 (1.25)	-1.09 (1.29)	-1.35 (1.55)	-0.87 (1.35)	-1.16 (1.62)
Motorola	1.76 (0.88) *	1.36 (0.85)	1.48 (0.93)	1.24 (0.9)	1.26 (1.02)	0.9 (1)	1.12 (1.11)	0.87 (1.06)
IBM	-1.5 (1.55)	-1.69 (1.54)	-1.38 (1.58)	-1.51 (1.58)	-1.12 (1.5)	-1.27 (1.49)	-1.07 (1.53)	-1.13 (1.5)
Time Trend	-3 (0.81) ***	-2.58 (0.77) ***	-3.43 (0.77) ***	-3.06 (0.72) ***				
Time Trend Sq.	0.07 (0.04) *	0.05 -0.04	0.09 (0.04) **	0.07 (0.04) *				
Intercept	84.52 (5.10) ***	82.7 (4.93) ***	87.74 (4.84) ***	86.1 (4.59) ***	77.56 (4.98) ***	75.5 (4.81) ***	80.18 (4.81) ***	78.01 (4.58) ***
Class Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Fixed Effects	-	-	-	-	Yes	Yes	Yes	Yes
R-Squared	60%	59%	58%	57%	63%	63%	62%	61%

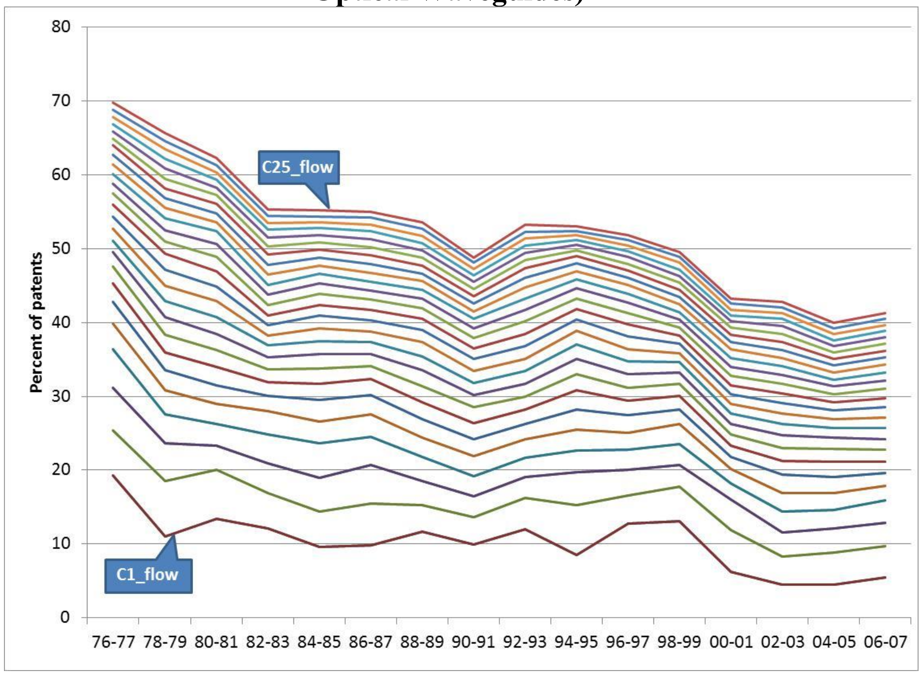
Notes: Regressions are ordinary least squares, with S.E. in parentheses. ***, **, * indicate significance at the 1%, 5%, and 10% levels respectively. Standard errors are clustered by class. An observation is a patent technology class and a two-year time period. N is 450 in odd numbered models, and 443 in even numbered models. Each model includes technology class fixed effects. Models 1-4 include a linear and a quadratic time trend; models 5-8 include time fixed effects. The sample includes the highest quartile of patents in the period 1976 to 2007 that are ultimately granted by the USPTO on or before 2010, where quality is measured by citations received.

Figure 1: Granted Patents by Application Year



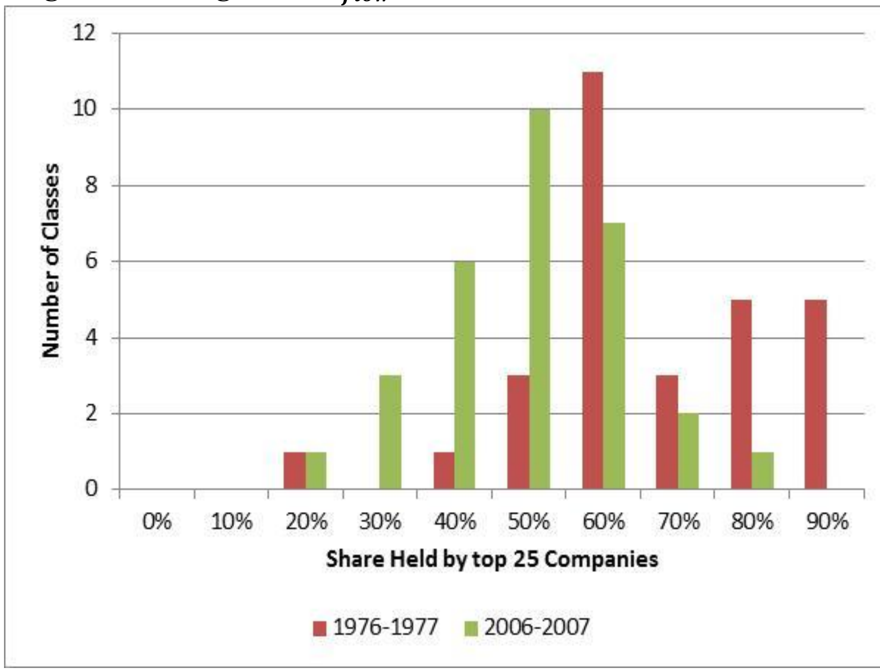
Notes: The sample includes patent applications from thirty patent technology classes in the ICT equipment industry from 1976 to 2007 that are ultimately granted on or before 2010, at all levels of patent quality.

Figure 2: Patent Flow Concentration Levels (Technology Class 385, Optical Waveguides)



Notes: The sample includes patent applications from the Optical Waveguides technology class (class 385) from 1976 to 2007 that are ultimately granted on or before 2010, at all levels of patent quality. The years are grouped into two year cells. The concentration is measured by the share of top *i* firms in terms of patent applications within each two-year cell, where *i* ranges from 1 to 25.

Figure 3: Change in $C25_{flow}$ distribution from 1976-77 to 2006-07



Notes: The sample includes patent applications from the thirty ICT equipment industry patent technology classes from 1976 to 2007 that are ultimately granted on or before 2010, at all levels of patent quality. The concentration is measured by the share of top twenty-five firms in terms of patent applications within each two-year cell.

Appendix I. Construction of Patent Data

This Study uses patent data constructed from raw USPTO text files for the period from 1976 to 2010 for a variety of reasons. First, coverage of the NBER data files ends in 1999 for the inventor variables, and in 2006 for the remainder of the data; our newly constructed data set goes to 2010. In addition, the NBER data do not include the original names of patent assignees; instead the data provide assignee names that have gone through a series of standardizations. We use the original names from the newly constructed data in the process of linking the patent data to the M&A data as described below.

Each week the USPTO makes available a new XML file, which can be accessed on its FTP site, containing bibliographic information for the patents granted within the prior week. In addition, the USPTO makes historical files available through the Google Patents Bulk Downloads site. This Study supplements the NBER patent data period with the XML files that go back to 2001, and the yearly hierarchical text files that cover the 1976-2001 period, resulting in the utilization of 474 weekly XML files and twenty-six yearly text files.[91] The newly organized data include information on granted utility patents applied for and granted between 1976 and 2010, including the application year, grant year, patent technology class, patent assignee name, location, and type.

In order to verify the data quality, we conduct extensive comparison of the newly compiled data against NBER patent data files for the overlapping period. In addition, we compare various aggregate statistics against the USPTO aggregate patent statistics. Table A1 presents patent counts by grant year from our data and the USPTO aggregate statistics page. As observed in the Table, the two datasets follow each other very closely. Comparisons on other patent properties follow similarly close trends.

In addition to the main bibliographic items, the USPTO assigns a primary technology class and a number of secondary technology classes to each patent at the time of grant. The classification system may be modified over time due to advances in technologies or other reasons. The USPTO updates the technology classes of all patents granted since 1790

91. Between 1976 and 2010 the data format changed dramatically, once in 2002 and again in 2005. Some minor changes were also made in 2006. The corresponding variables from various years were matched using the relevant version of the Redbook documentation from the USPTO website.

and publishes them in the MCF once every two months. Our data include classifications from the December 2010 version of this product.

As in prior work, we take advantage of citations. The patent data contain the citations made by the granted patents between 1976 and 2010 to other granted patents in earlier periods. This information is used in controlling for the heterogeneity in patent value, which has a highly skewed distribution.[92] Prior studies have documented a strong, positive correlation between the value of a patent and the number of citations it receives.[93] In keeping with this literature, we control for the quality of patents and repeat the analyses on the sample of highly cited patents, in addition to conducting our analyses on the entire sample of granted patents.

The main pillar of this Study is the patent ownership composition, which is constructed using the share of granted patents to each unique assignee. However, the newly compiled USPTO patent data does not contain a unique assignee identifier (akin to NBER's pdpass variable) that is consistent across different patents and across time. The main assignee identifier is the firm name, which is a long string and is susceptible to errors in links due to potential misspellings, different spelling of foreign firms, and differences in abbreviations. To address the lack of unique firm identifiers, we developed a methodology to link different name strings representing the same entity to each other. We discuss the details of this algorithm and a comparison to NBER's unique identifiers in Appendix II.

---

92. Harhoff et al., *supra* note 70, at 512 (discussing the highly skewed distribution of values); Pakes & Schankerman, *supra* note 70, at 86 (same).

93. *See, e.g.*,. Harhoff et al., *supra* note 70.

Table A1: Granted Utility Patents

Grant Year	USPTO	XML Compilation	Difference
2010	219,614	219,909	295
2009	167,349	167,553	204
2008	157,772	157,894	122
2007	157,282	157,502	220
2006	173,772	173,922	150
2005	143,806	143,927	121
2004	164,290	164,413	123
2003	169,023	169,104	81
2002	167,330	167,424	94
2001	166,035	166,158	123
2000	157,494	157,595	101
1999	153,485	153,592	107
1998	147,517	147,576	59
1997	111,984	112,019	35
1996	109,645	109,653	8
1995	101,419	101,431	12
1994	101,676	101,696	20
1993	98,342	98,384	42
1992	97,444	97,473	29
1991	96,511	96,557	46
1990	90,365	90,421	56
1989	95,537	95,566	29
1988	77,924	77,937	13
1987	82,952	82,967	15
1986	70,860	70,865	5
1985	71,661	71,669	8
1984	67,200	67,215	15
1983	56,860	56,860	0
1982	57,888	57,878	10
1981	65,771	65,766	5
1980	61,819	61,812	7
1979	48,854	48,839	15
1978	66,102	66,084	18
1977	65,269	65,200	69
1976	70,226	70,190	36
Total	3,911,078	3,913,051	2,293

Notes: Patent counts by grant year from the USPTO aggregate patent statistics and our newly constructed sample from USPTO XML and text files. Source: U.S. Patent Statistics Chart, Patent Technology Monitoring Team (PTMT), USPTO. Last accessed February 22, 2012, http://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm.

Appendix II: Firm Name Linking Algorithm

The newly compiled USPTO patent data do not contain a unique assignee identifier that is consistent across time. The main assignee identifier is the firm name, which is a long string and is susceptible to errors in links due to potential misspellings, different spelling of foreign firms, and differences in abbreviations (such as "corporation," "co.," and "co"). To address the lack of unique firm identifiers, we developed a methodology to link different name strings representing the same entity to each other.[94]

The linking algorithm consists of two stages: an automated stage and a human intelligence stage. In the automated stage, a computer program standardizes the firm names using common abbreviations and misspellings identified from the data, such as "corp," "corporation," "corp<u>oo</u>ration," etc. The program then conducts a linking based on common words in company names. Although this program captures a significant portion of actual matches across datasets, it also produces false positives. An example of a false positive would be flagging "ABC Business Solutions" and "XYZ Business Solutions" as the same company due to the common "Business Solutions" phrase. To work around this problem we conduct a human intelligence stage. In this stage the matches identified by the computer program are fed into a crowd-sourcing website, Amazon's Mechanical Turk, for manual human verification that will un-flag the false positives, and leave only the actual matches for use in the data linking. [95]

As a quality check of this process we compare the results to the NBER patent data files, which address the same issue only within the patent data, and mapped 322,783 names into 243,800 unique entities. A comparison of the results from our algorithm on a sample of 70,000 firm names to the NBER patent data file suggests that our results are as good as the NBER matches, if not better.

Differences exist between the two algorithms, partly due to random errors and partly due to the difference in what is considered a unique entity. Table A2 provides an illustration through a subset of names for

94. This new variable will assume the role of the NBER patent data's pdpass variable in our dataset.

95. Crowdsourcing sites enable the outsourcing of simple tasks to a large group of workers on demand. In our case, workers see a pair of company names matched by the computer program, and are asked to simply choose "yes" or "no" to indicate whether the two companies are the same companies, or not. Outsourcing the linking process to a large workforce and using standard quality control techniques facilitates the timely completion of the task at a reasonable cost.

the Sony Corporation. In this list, each line represents a different entity (different pdpass) in the NBER data, whereas all are considered part of the same entity in our data. The three versions of "Sony Electronics Inc." being assigned to different entities in the NBER data give an example of random errors in the matching process.[96] However, designating "Sony Corp of America" and "Sony Electronics Inc." as different entities highlights differences in what we consider a firm. In this assignment we believe that firms create different subsidiaries for a variety of reasons, including tax blueprint, legacy, and other managerial or strategic issues. However, we conjecture that two such firms would go through patent infringement issues only under very extreme, unlikely conditions; therefore we consider them the same entity.

96. Similar cases where a match missed by our algorithm is captured by the NBER also exist in the data. Table A2 does not indicate superiority of our algorithm over NBER's.

Table A2: Assignee Names for SONY Corp.

NBER pdpass	NBER Assignee Name
11297047	SONY AUSTRALIA PTY LTD
11277610	SONY BROADCAST & COMMUNICATION
11958546	SONY CHEM CORP
13040458	SONY CHEM CORP NEAGARI PLANT
12059716	SONY CINEMA PROD CORP
12104210	SONY COMPUTER ENTERTAIMENT INC
12805945	SONY COMPUTER ENTERTAINMENT AM
13147302	SONY CORP ENTERTAINMENT AMERIC
11205194	SONY CORP OF AMERICA
13171917	SONY CORPORATIOM
21878152	SONY ELECTONICS INC
21589106	SONY ELECTRONIC INC
11399266	SONY ELECTRONICS INC

Appendix III: ICT Equipment Patent Technology Classes Considered

Class	Description
178	Telegraphy
330	Amplifiers
331	Oscillators
332	Modulators
333	Wave transmission lines and networks
334	Tuners
340	Communications: electrical
342	Communications: directive radio wave systems and devices (e.g., radar, radio navigation)
343	Communications: radio wave antennas
348	Television
358	Facsimile and static presentation processing
367	Communications, electrical: acoustic wave systems and devices
370	Multiplex communications
371	Error Detection/Correction and Fault Detection/Recovery
375	Pulse or digital communications
379	Telephonic communications
380	Cryptography, subclasses 255 through 276 for a communication system using cryptography
381	Electrical Audio Signal Processing Systems and Devices, subclasses 1+ for broadcast or multiplex stereo
385	Optical waveguides
398	Optical communications
455	Telecommunications
700	Data processing: generic control systems or specific applications
701	Data processing: vehicles, navigation, and relative location
702	Data processing: measuring, calibrating, or testing
703	Data processing: structural design, modeling, simulation, and emulation
704	Data processing: speech signal processing, linguistics, language translation, and audio compression/decompression
705	Data processing: financial, business practice, management, or cost/price determination
706	Data processing: artificial intelligence
707	Data processing: database and file management or data structures
708	Electrical computers: arithmetic processing and calculating
709	Electrical computers and digital processing systems: multicomputer data transferring
710	Electrical computers and digital data processing systems: input/output
711	Electrical computers and digital processing systems: memory
712	Electrical computers and digital processing systems: processing architectures and instruction processing (e.g., processors)

713	Electrical computers and digital processing systems: support
714	Error detection/correction and fault detection/recovery
715	Data processing: presentation processing of document, operator interface processing, and screen saver display processing
716	Data processing: design and analysis of circuit or semiconductor mask
717	Data processing: software development, installation, and management
718	Electrical computers and digital processing systems: virtual machine task or process management or task management/control
719	Electrical computers and digital processing systems: interprogram communication or interprocess communication (ipc)
720	Dynamic optical information storage or retrieval
725	Interactive video distribution systems
726	Information security