Industry Location Shift through Technological Change - A Study of the US Semiconductor Industry (1947-1987)

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Engineering and Public Policy

Jonathan D. Kowalski

B.S., Electrical Engineering, Kettering University

M.S., Engineering and Public Policy, Carnegie Mellon University

Carnegie Mellon University

Pittsburgh, PA

May, 2012

Acknowledgements

This research was supported in part by grants from the National Science Foundation (SBE-0965451, "Clusters, Heritage and the Microfoundations of Spillovers - Lessons from Semi-Conductors"; SBE-0738182, "Evaluation of Research Groups: An Endogenous Approach"; SES-0727000, "Collaborative Research: Sustainability in Disruptive Technology Contexts: Dynamic Capabilities and Locus of Innovation"), the Richard King Mellon Foundation, the Claire and John Bertucci Fellowship, a Dean's Fellowship from the Carnegie Institute of Technology, the Engineering & Technology Innovation Management (ETIM) program at Carnegie Mellon and a research travel grant from the William P. Clements Center for Southwest Studies and the DeGolyer Library at Southern Methodist University. The views expressed in this dissertation do not necessarily reflect the views of these organizations.

Many individuals contributed to the success of this work. I would particularly like to thank my co-advisor Steven Klepper and my co-advisor and committee chair Francisco Veloso for helping me find a topic that I wanted to investigate and giving me the patient advice and tools to do so. I am also grateful for the feedback from the other members of my committee, Serguey Braguinsky and David Hounshell.

Other faculty and staff members at Carnegie Mellon have helped to inspire me and provide me with the guidance and support necessary to complete this dissertation and the Doctoral program, including Eden Fisher, Indira Nair, Suzie Laurich-McIntyre, Vicki Finney, Meryl Sustarsic, Paige Houser, Patti Steranchak and Barbara Bugosh.

The data-intensive nature of this dissertation and the novel data set created as a result of it could not have been completed without significant assistance from individuals at many different institutions. I would like to thank the library staff and inter-library loan (ILL) staff at Carnegie Mellon for helping me with countless requests, especially Joan Stein and Andrew Marshall, as well as the library and ILL staffs at the University of Michigan, Michigan State University, Lock Haven University, The Pennsylvania State University, the James J. Hill Reference Library and Texas A&M University.

While a large portion of the data required for this dissertation was able to be collected through inter-library loan, some travel was required as some of the sources that I used are considered rare. I would like to thank Rosa Summers, Scott Porter and Michelle Nyland for hosting me during my trips to the Minneapolis Public Library, Lock Haven University and Southern Methodist University, respectively. Additionally, I would like to thank the staff of the DeGolyer Library and William P. Clements Center for Southwest Studies for their assistance and hospitality during my visit to Southern Methodist University, especially Andrew Graybill, Andrea Boardman, Russell L. Martin III, Ada Negarayu and Pamalla Anderson.

Several individuals assisted me with processing the data once it was collected in order to create a single, useable dataset, including Jeff Plotzke and my research assistants Xing Xin, Joel Feinstein and Ben Streeter.

While data work accounted for a large portion of my time gathering information for the dissertation, I was also fortunate to spend some time talking to several individuals who helped to develop semiconductor technology and the industry which I have come to know so well. To

Barry Bebb, Harvey Cragon and Ross Macdonald, thank you for taking the time to discuss your experience at Texas Instruments and for your honesty and openness. To Ed Millis, thank you for taking the time to communicate with me and for the copy of your book.

This dissertation is the product of my own writing and ideas, but also resulted from input from many others. I would like to extend thanks to Francisco Veloso, Steven Klepper and David Hounshell for providing extensive development feedback and edits throughout the writing process, as well as my proofreaders, Sharad Oberoi and Kerry Jo Richards.

Many other graduate students, past and present, have helped me with my research and made life enjoyable – both here in Pittsburgh and throughout the United Stated - including Anu Narayanan, Adam Tagert, Carolyn Denomme, Ruth Poproski, Natalie Scala, Austin Mitchell, Pete Versteeg, Eyiwunmi Akinsanmi, Mike Blackhurst, Timothee Doutriaux, Carla Costa, Rodrigo Belo, Ana Venâncio, Cristina Carias, Xu Li, Andreia Rafael, Patrick Gage Kelley, PJ Dillon, Debra Jo Scardino, Denise Philpot, Matt Cooper, Alex Evans, Kevin McComber, Donna Dueker, and Jason Heustis.

My success at Carnegie Mellon was in a large part due to my prior educational endeavors, and I would like to thank the faculty, staff and my friends at Kettering University who provided me with incredible educational opportunities. I would like to specifically thank Laura Sullivan for pushing me to think outside of the box and be daring in my pursuits, Karen Palmer for encouraging me to pursue graduate school while serving as my undergraduate thesis advisor, Donna and Mark Wicks for serving as a family away from home and pushing me to always

achieve more, and the entire Business Department, especially Andy Borchers and Dave Strubler, for providing presentation opportunities for my research during my time at Carnegie Mellon.

Finally, I am who I am because of my family. My time at Carnegie Mellon, while rewarding and enriching, was not without its challenges. My family was always there to support me, cheer me on, listen, and inspire me to pursue my passions. I would like to thank my parents, brothers and extended family for their unconditional love and support throughout my graduate studies. I can always count on my family, both immediate and extended, for support, and for that I am eternally grateful.

Abstract

Silicon Valley is a storied region regarded by many as a model for economic development. Many governments have attempted or considered implementing policies or projects aimed at re-creating the success of Silicon Valley. However, it is not clear that we truly know what led to Silicon Valley's success, as existing work has not pursued industry-wide firm-level analyses to examine the mechanisms that allowed Silicon Valley to emerge as a key region. This work seeks to begin to address this literature gap in order to better inform regional economic development policy moving forward. In examining the development of Silicon Valley and the semiconductor industry, a detailed analysis of the technological developments leading to both transistors and integrated circuits was performed. From this analysis, it became clear that the nature and availability of knowledge changed significantly between the transistor and integrated circuit eras, with knowledge becoming more complex, tacit, and less available throughout the industry. From this understanding, specific predictions and hypotheses regarding firm and industry development were generated guided by existing theory.

Using a novel dataset of US semiconductor production between 1947 and 1987, this dissertation examines empirically the development of the semiconductor industry to test these hypotheses. The results show that the mechanisms driving success differed between the two eras of the semiconductor industry. As the industry transitioned to the transistor era, existing electronics firms dominated the industry, which resulted in a build-up of transistor firms in the same clusters that previously produced electronics products; however, this was not the case as the industry transitioned to integrated circuits. The nature of the knowledge in the integrated circuit era

allowed spinoff firms to emerge as an important force in the industry, out-performing incumbent firms, which ultimately led to the emergence of Silicon Valley as the primary semiconductor industry cluster. It is important to understand the technological context that created an opportunity for spinoff firms to fuel Silicon Valley's ascension to significance within the industry, as this dissertation demonstrates that the applicability of existing theories regarding firm entry and development are influenced by the nature of technology. Understanding the conditions under which various mechanisms can be effective in promoting firm entry and performance, and thus regional clusters is vital in order to craft efficient public policy and projects aimed at building industry clusters in the future. This dissertation contributes greatly to that understanding.

Table of Contents

Acknowledgements iii

Abstract vii

Table of Contents ix

List of Figures xii

List of Tables xiv

Chapter 1: Introduction 1

Chapter 2: Theory 7

Geographic Advantage 7

Firm Heritage 10

Similar Results, Different Mechanisms 15

Chapter 3: The Evolution of Semiconductor-based Electronics 18

From Cat's Whisker to Transistor 19

Transistor Knowledge Access through
Government Intervention and Competitive Interests 29

Hypotheses: Transistor Market Entry and Performance 36

From the Point Contact Transistor to Integrated Circuits 47

Stepping Back 155

Chapter 4: Research Data 159

Firm Identification and Production Data 159

Performance Data 163

Heritage Data 164

Chapter 5: Transistor Entry & Performance 169

Transistor Producer Location Distribution 169

Transistor Producer Background Distribution 171

Firm Entry Analysis 172

Firm Performance Analysis 177

Survival Analysis 179

Market Share Attainment 183

Chapter 6: Integrated Circuit Entry & Performance 189

Integrated Circuit Producer Location Distribution 189

Integrated Circuit Producer Background Distribution 191

Firm Entry Analysis 193

Firm Performance Analysis 196

Survival Analysis 198

Market Share Attainment Analysis 202

Chapter 7: Organizational Learning and Adaptation in the Semiconductor Industry 207

Technological Frontier at Entry 207

Adoption of Technological Frontier Following Entry 213

Spinoff Learning through Experience Gained at Parent 217

Spinoff Performance and Influence of Technological Frontier Production 226

Chapter 8: Conclusions & Policy Implications 240

References 245

Appendix A: CMSA Definition Maps 257

Appendix B: Product Categories 259

List of Figures

Figure 1: Cat's Whisker Rectifier. 20

Figure 2: Vacuum Tube (Triode) 22

Figure 3: First Point Contact Transistor. 29

Figure 4: Teal's Crystal Growing Apparatus. 50

Figure 5: Grown Junction Transistor. 52

Figure 6: Alloy Junction Transistor. 54

Figure 7: Mesa Transistor. 72

Figure 8: Planar Transistor. 86

Figure 9: Sub-Assemblies of the RCA Micro Module. 100

Figure 10: First Integrated Circuit, Designed by Jack Kilby at Texas Instruments. 116

Figure 11: First Planar Integrated Circuit, Designed by Jean Hoerni at Fairchild. 120

Figure 12: Semiconductor Network Manufacturing Yields 125

Figure 13: Planar Production Process Diagram. 131

Figure 14: Equipment Used for Semiconductor Diffusion. 133

Figure 15: Transistor Firm Location over Time 170

Figure 16: IC Firm Location over Time 190

Figure 17: 5-Year Moving Average Share of Entrants by IC Type..... 208

List of Tables

Table 1: Comparison of Major Transistor Types, 1947-1962. 92

Table 2: Firms pursuing circuit miniaturization in 1961 97

Table 3: Major steps required for production of monolithic integrated circuits and sources of information for each step. 123

Table 4: Transistor Firms by Background and Region, 1949-1987..... 172

Table 5: Transistor Entry Analysis 175

Table 6: Transistor Survival Analysis..... 181

Table 7: Transistor Market Share Attainment Analysis 185

Table 8: Summary of Transistor Results 188

Table 9: Integrated Circuit Firms by Background and Region, 1965-1987. 192

Table 10: Integrated Circuit Entry Analysis 195

Table 11: Integrated Circuit Survival Analysis 200

Table 12: Integrated Circuit Market Share Attainment Analysis 203

Table 13: Summary of Integrated Circuit Results 206

Table 14: Integrated Circuit Firms by Background and Region, 1965-1987. 210

Table 15: Monolithic at Entry Analysis 211

Table 16: Adoption of Monolithic ICs Following Entry 215

Table 17: Expanded Monolithic at Entry Analysis 218

Table 18: Broad Integrated Circuit Product Categories 220

Table 19: Likelihood (Probit) that Spinoff Produces a Product Given that Parent Produced It - Aggregate Product Categories 222

Table 20: Likelihood (Probit) that Spinoff Produces a Product Given that Parent Produced It - Disaggregated Product Categories 224

Table 21: Share of Integrated Circuit product categories produced by parents and spinoffs 225

Table 22: Expanded Transistor Survival Analysis 230

Table 23: Expanded Transistor Market Share Attainment Analysis 231

Table 24: Expanded Integrated Circuit Survival Analysis 235

Table 25: Expanded Integrated Circuit Market Share Attainment Analysis 236

Table 26: Summary of Organizational Learning and Adaptation Results 237

Table 27: Broad Integrated Circuit Product Categories 259

Table 28: Narrow Integrated Circuit Product Categories 259

Chapter 1: Introduction

Today, Silicon Valley is a storied region regarded by many as a model for economic development. This attention is certainly warranted. While some electronics production existed within the region in the beginning in the 1910's,[1] it was during the second half of the 20th century that Silicon Valley developed into one of the most influential centers of high-tech innovation. This development was driven by the electronics industry – specifically manufacturers of integrated circuits. The relatively young Silicon Valley firms came to dominate the semiconductor industry. By 1985, eight Silicon Valley firms alone accounted for approximately forty-nine percent of the semiconductor industry's output (Klepper, 2009a) – quite a feat for any region. The region became so important in the context of the industry that the material used in the manufacture of semiconductors was adopted as the region's name. There seems to be little doubt that semiconductors drove Silicon Valley's emergence as an economic center, as the population of Santa Clara county, the heart of Silicon Valley, quadrupled over the first 30 years of the industry (Klepper, 2010, p. 15) and high technology employment in the Valley increased by 900% during a similar period (Saxenian, 1994, p. 3).

[1] See Sturgeon (2000) for a discussion about De Forest's invention of the vacuum tube in the Silicon Valley region as well as the discussion in Morgan (1967) regarding early electronics activities there.

Since the 1960's, regions and countries around the world have discussed how to replicate the success of Silicon Valley (Hamilton, 2008; Leslie & Kargon, 1996).[2] In recalling policy recommendations regarding such replication, Moore and Davis eloquently describe the typical approach, which they call the "magic potion":

> *"Combine liberal amounts of*
>
> *Technology*
>
> *Entrepreneurs*
>
> *Capital, and*
>
> *Sunshine.*
>
> *Add one (1)*
>
> *University.*
>
> *Stir vigorously.* (Moore & Davis, 2004, p. 9)*"*

The Skolkovo Innovation Center project launched on the western outskirts of Moscow is one of many attempts to replicate Silicon Valley following this "magic potion" very closely (perhaps with the exception of the sunshine).[3] Russia's President Medvedev recently visited Silicon Valley in an effort to learn more about the region that he hopes to emulate in Skolkovo, noting

---

[2] While a number of efforts exist, the two references refer to projects in Colorado, Indiana, Michigan, Pennsylvania, Illinois, Tennessee, North Carolina, South Carolina, New Jersey, Texas and South Korea.

[3] The weather metric for days with sunshine appears to be clear days ("Climate of Moscow," n.d.). San Jose, the "capital" of Silicon Valley has more than 300 sunny days per year ("City of San Jose," n.d.), while Moscow only has 82 days with sunshine ("Climate of Moscow," n.d.), clearly placing Moscow at a disadvantage with respect to the Silicon Valley area.

that “We [Russia] have money, but we don’t have Silicon Valley(“Why Dmitry Medvedev wants a Russian Silicon Valley,” n.d.).” The project aims to “concentrate international intellectual capital, thereby stimulating the development of break-through projects and technologies” within Russia. A massive undertaking (“What is Skolkovo?,” n.d.), the Skolkovo Innovation Center is a city that is being built from scratch, complete with industrial technology parks, universities, housing and infrastructure (“What is Skolkovo?,” n.d.) in an effort to concentrate people, ideas, and firms in a single region. The investment doesn’t end with physical buildings, however, as the Russian government has pledged tax incentives, grants and financing for entrepreneurs, and other benefits for individuals interested in pursuing innovative ideas at Skolkovo (“Skolkovo Participant Benefits,” n.d.).

While the Skolkovo project may be one of the boldest attempts to replicate Silicon Valley, interest in doing so has been widespread in the last 50 years (Hamilton, 2008; Leslie & Kargon, 1996). In fact, many governments have taken interest in replicating the success of Silicon Valley, as the area is one of the most notable examples of a region that developed into an industrial cluster. Clusters can exist within regions, but clusters and regions are not synonymous. Rather, “clusters are groups of interconnected firms and industries in the same field that arise in particular economic areas” (Porter, 2001, p. 140) and are typically associated with higher productivity, a high intensity of innovative activity, and new business formation (Porter, 2001, p. 140). If these are the results of a cluster emerging within a region, it seems that governments throughout the world would be interested in cluster development.

The European Union has placed a strong emphasis on the use of resources toward the development of regional clusters. To help "maintain a permanent policy dialogue at [the] EU level among national and regional authorities responsible for developing cluster policies and managing cluster programmes in their countries," the European Cluster Alliance was formed in 2006 ("European Cluster Alliance," n.d.). With over 110 members from across the European Union, the group is only one indication of the importance placed on cluster development within the EU. While aggregate data on government expenditures for cluster development are not available, approximately $55 million was spent on cluster projects in the Czech Republic alone during a 10-year period (2004-2013) (Skokan, 2011). This amount, however, pales in comparison to the cost of an initiative like Skolkovo. Estimates from the Skolkovo Foundation indicate that expenditures for the innovation center will reach approximately $5.7 billion over six years (2010-2015) (*Skolkovo Innovation Center: Executive Summary*, 2010).

Given the amount of attention and resources that are being devoted to cluster development throughout the world, it seems vital to have a clear understanding of the mechanisms that allowed Silicon Valley to emerge and evolve to its current state. It is only with this understanding that more effective public policy aimed at regional development can be crafted. The fascination with the development of Silicon Valley has produced many qualitative studies (Kaplan, 2000; Lécuyer, 2006; Malone, 1985; Saxenian, 1994) reporting the region's intriguing story and advancing explanations for its success. Therefore, the stylized story is well known. It began with the invention of the transistor at Bell Labs in Murray Hill, NJ in 1947 by William Shockley, who won the Nobel Prize for his invention. Following this development, Shockley left Bell Labs to start his own company, which he located in Mountain View near his hometown

of Palo Alto, CA. Shockley was a magnet for talent and recruited many outstanding people from around the country, but he was also a poor manager. Poor treatment of his employees and inconsistent leadership led many of the early top recruits to leave and form Fairchild Semiconductor (Lécuyer, 2006). Fairchild went on to invent a key technology that led to the development of the integrated circuit – the planar process (Lécuyer, 2006). The interaction of firms with Stanford University, along with the pleasant weather and the flexibility of the firms and culture of cooperation within the Valley, seem to have created the perfect combination for the creation of a high-tech cluster (Saxenian, 1994). Over time, Silicon Valley would account for a majority of the semiconductor sales in the United States, with the region becoming the cluster for high-tech innovation. These accounts, along with other anecdotal evidence, have provided a strong basis for an impression that the co-location of unique firms, talent and institutions fueled Silicon Valley's growth.

While much attention has been paid to the development of Silicon Valley, there has been little quantitative work examining the growth of the semiconductor industry within the valley and also across the rest of the US.[4] Over time, the story of Silicon Valley has been retold (Kaplan, 2000; Malone, 1985; Saxenian, 1994) so often that it now seems to be true, with little examination of detailed empirical evidence at the firm-level. While it was clear that a firm-level analysis would

[4] Quantitative work, however, does exist with respect to worker mobility and its effect on knowledge transfer within regions such as Silicon Valley (Almeida & Kogut, 1999; Fallick et al., 2006) and the emergence of inventor networks within the region (Fleming & Frenken, 2007). While a great deal of work exists on the region (Bresnahan & Gambardella, 2004; Kaplan, 2000; Lécuyer, 2006; Malone, 1985; Moore & Davis, 2004; Saxenian, 1994), most of this work focuses on a single firm, is done at the regional level, or provides a historical narrative regarding the region as opposed to a firm-level examination of the evolution of the region and semiconductor industry.

be a welcome contribution to the literature, it was realized that any analysis needed to also consider the technological context of the industry. In order to form specific hypotheses from the general theories regarding firm and industry development, a detailed analysis of semiconductor technology during the transistor and integrated circuit eras has been pursued, revealing changes in the technological knowledge required to operate at the technological frontier of the industry between the eras. Using the hypotheses generated, a novel dataset of semiconductor producers has been created which provides a unique opportunity to examine empirically the early semiconductor industry and to test hypotheses regarding firm and industry development.

This dissertation is structured as follows. Chapter 1 serves as an introduction to the dissertation, providing motivation for examination of the topic. Chapter 2 will discuss prevalent theories of firm entry and performance based on geographic advantage as well as firm heritage and capabilities. Chapter 3 will detail the history of the technology in the semiconductor industry and relate the technology and industry context to the theories presented in the previous chapter to generate specific hypotheses. The research data and the evolution of the semiconductor industry will be discussed in more detail in Chapter 4. Chapters 5 and 6 will examine how firm attributes affected firm entry into transistors and integrated circuits, respectively, and firm performance within these products following entry. The interaction of firms with the technological frontier - both at entry as well as following entry - will be examined in Chapter 7, as well as whether knowledge regarding products is transferred to spinoff organizations from parent firms. Finally, the thesis will conclude by drawing conclusions from the research and describing relevant policy implications.

Chapter 2: Theory

To examine why Silicon Valley was able to emerge as a notable semiconductor region, this dissertation begins by looking at existing theories that help us understand this important phenomenon. This chapter reviews two leading theories describing regional development and explores their predictions regarding firm entry and performance.

Geographic Advantage

One explanation for the development of regions that come to dominate specific industries (in terms of employment, production volumes, innovations, etc.) is based on geography. This theory is rooted in the idea that being co-located with other firms offers advantages to firms and individuals, generating agglomeration economies. The existence of agglomeration economies in the literature dates back to the late 19th Century, when Alfred Marshall first described three potential sources of agglomeration economies (Marshall, 1890). These have been further refined in more recent work (Fujita, Krugman, & Venables, 2001; J. Henderson, 1998; J. V. Henderson, 1974) and can essentially be described as:

- the sharing of inputs (both raw materials and specialized inputs, which can decrease the cost of such inputs due to factors such as scale economies or logistics)
- labor market pooling (allowing for better matches of employees and employers)
- knowledge spillovers (allowing firms to have easier and better access to knowledge from other local firms)

If advantages for production exist within regional clusters, firms located in these regions are expected to out-perform other firms that do not enjoy these advantages due to their location outside of industry clusters. In fact, previous work shows that firms in clusters are more

innovative (Baptista & Swann, 1998) and more productive (J. Henderson, 2003), lending support to this idea. The expectation can be summarized by the following statement:

> *Firms located within existing clusters are more likely to out-perform firms located outside of clusters following entry, all else held equal.*

Over the last several decades, knowledge spillovers have gained particular importance in the context of the agglomeration literature, especially in knowledge-based industries. It has been observed that these industries tend to cluster geographically, and a variety of studies suggest that firms may seek out knowledge spillovers by locating near other firms with the aim of taking advantage of them (Audretsch & Feldman, 1996; Carlton, 1983; Chung & Alcacer, 2002; Head, Ries, & Swenson, 1995).

Two specific mechanisms can help explain knowledge spillovers: one is the informal exchange of information and the second is the mobility of employees across firms. Von Hippel (1987) and Bassett (2002) document how engineers routinely exchange valuable proprietary technological information with other firms in their industry, including rivals. Furthermore, when knowledge is sufficiently tacit and embodied in human capital, technological spillovers may require the movement of an employee from one firm to another, rather than just informal contacts between employees.

A cluster might contribute to the establishment of spillovers through both mechanisms. First, agglomeration facilitates the process of identifying and meeting individuals who might be valuable sources of knowledge for a firm by maximizing mutual accessibility, thus facilitating knowledge exchange (Duranton & Puga, 2004; Helsley & Strange, 1990). Second, if relocation

is costly and thus employees with critical knowledge are likely to stay local, the presence of a cluster will allow greater mobility and as a result, facilitate diffusion of relevant knowledge in the cluster, helping the firms located there (D. Cooper, 2001; Pakes & Nitzan, 1983). Empirical research has provided some support for these theories. For example, Almeida and Kogut (1999) as well as Fallick et al. (2006) show that inventors and knowledge workers in semiconductors and computers, respectively, are indeed more mobile in clusters. Breschi and Lissoni (2006, 2009) show that inventors typically move within the region where they are located and that such mobility can largely explain higher rates of citations across co-located firms. The establishment of a cluster itself may result from spillovers. If hiring knowledgeable employees from other firms is critical for innovation and profitability, and relocation to a different region is sufficiently costly, firms may be better off clustering (Combes & Duranton, 2006; Fosfuri & Ronde, 2004), especially given rapid technological progress. Yet, prior work has mostly focused on patents and citations and has not explored more directly the consequences of spillovers for firm product choices and performance.

If clustering helps knowledge diffusion through greater informal contacts and inter-firm mobility, we should expect existing firms located in a cluster to learn about the technological frontier and to be able to adopt it sooner than firms located elsewhere. More specifically, we should expect that:

> *The rate at which existing firms not producing at the technological frontier subsequently adopt products at the technological frontier at any given moment is greater for producers located in clusters.*

If these predictions are true, agglomeration theory can help describe how a successful cluster would form. Once several firms producing a given product emerge within a region, it would become advantageous to operate within that region producing the same (or related) products, which would fuel growth of the industry within the region. This growth may come from existing producers that choose to enter the product area with an advantage, new firms that may be founded in the region, and even existing producers from other regions that choose to relocate to access the regional advantages. Additionally, when technological change occurs, firms within this region will be fastest to adopt the technological change, leading to an advantage with respect to operating at the technological frontier. These advantages allow firms in the region to out-perform firms elsewhere, leading to higher survival rates and performance, a higher proportion of firms operating at the technological frontier, and in general a prominence within the industry. In many cases, this story describes the development of Silicon Valley in the way that it has been portrayed.

Firm Heritage

While advantages derived from agglomeration economies can be used to describe and predict industry development within regions, a complementary perspective of how firms and regions can develop considers that firms require specific capabilities in order to begin producing specific types of products. These capabilities are gained through heritage, either through previous organizational experience, or through the experience of the organization's founder(s) in previous employment. This theory posits that knowledge may be tacit and gained only through

experience, and this experience is used to describe differences in both a firm's propensity to enter an industry as well as its performance following entry.

Broadly, firms that enter a new industry can be of three natures: diversifiers, startups, and spinoffs (Klepper, 2009b). Diversifiers are firms that have previously produced other products before beginning production in the context of a new industry. Startups are new firms that have not previously been in any market, founded by individuals who have no relevant experience related to the industry they are entering. Spinoffs are similar to startups as they are new firms, but individuals with previous related industrial experience found these. Previous experience consists of two components, each addressing one type of firm:

- For diversifying firms, the amount of previous technical experience producing related products and the relevance of such experience.
- For new firms, the previous experience of the firm's founder(s) at other firms.

The theory regarding these two components will be discussed separately in the following sections to motivate broad predictions that follow from the theory.

Capabilities Gained through Previous Experience

Any firm consists of people, equipment, processes, and institutional memory - all of which have the ability to shape future decisions. Existing literature suggests that a firm's previous technical experience conditions its propensity to enter related industries (Scott-Morton, 1999; Klepper & Simons, 2000; Lane, 1989) in part because the costs associated with entry (search, capital, development, etc.) are reduced due to the previous experience of the firm (Scott-Morton, 1999). Additionally, the same firms that are most likely to enter a new product type because of their

background have also been found to be the most successful, long-lived, and innovative among diversifying firms (Klepper & Simons, 2000; Lane, 1989). Television producers are a particularly salient example of this theory in action, as radio producers "entered earlier, survived longer, and had larger market shares than nonradio producers" (Klepper & Simons, 2000).

If a firm's related technical experience increases the likelihood that a firm will begin producing a product, the following should be observed:

> *Existing firms with relevant technical experience have a higher propensity to enter related products, all else held constant.*

Furthermore, if a firm's related technical experience contributes to a firm's performance, these firms should be more likely to out-perform other firms, summarized by the following statement:

> *Existing firms with relevant technical experience out-perform other firms (either new firms or firms without relevant technical experience) following entry, all else held constant.*

Spinoff Firms

While the heritage of a given firm seems to have influence over the firm's actions and performance, an organization's heritage doesn't end with the decisions that are made within the organization. In a similar fashion to the way that firms build capabilities and knowledge over time as a result of their production experience, the same is true for individuals who gain capabilities and knowledge over time through their own work experience. As people gain

experience and knowledge within existing firms, they can be driven to create their own companies, known as "spinoffs" (Klepper & Thompson, 2010, p. 2).

In a study examining the spinoff populations of the automotive and semiconductor industries,[5] Klepper (2009a) finds that the vast majority of spinoff firms were formed as a result of either managerial or strategic disagreements. An example of the type of disagreement that could lead to a spinoff is discussed by Fleming and Marx, where a software engineer employed by a firm had a dispute over the ownership of a recent business plan the engineer had developed. When the firm refused to grant the engineer ownership of the idea, the engineer stated, "Fine. I have a better idea for a company than anyone here and I'll go do it myself (Fleming & Marx, 2006, p. 18)".

This example demonstrates an important concept regarding spinoff firms, namely that employees are able to take the knowledge and expertise that they have developed through employment with them when leaving their employer, allowing for the transfer of valuable and critical knowledge to other firms. This situation is particularly troublesome for knowledge-based firms such as high technology firms, whose "most valuable assets 'walk out the door every night'" (M. Marx, Strumsky, & Fleming, 2009, p. 875). Employees with previous related experience who go on to found their own firms have the knowledge needed to compete in the industry and have the power to make the decisions that affect the new firm's strategy, thus allowing them to steer their

[5] These are two of many industries where a large proportion of firms have been identified as spinoffs. Some of the other industries include laser manufacturers, disk drive firms, biotech and Silicon Valley-based law firms (Klepper & Thompson, 2010, p. 2).

organization into decisions and markets based on the knowledge they gained during their previous employment. Examples of this occurring include two automotive companies, Brush Runabout and Oakland, which were both founded by former employees of Cadillac and became leading automotive firms (Klepper, 2007a). Empirical work finds that firms with this knowledge and power out-perform other new firms, and in many cases out-perform other firms with previous experience (Klepper, 2007a). However, the question remains as to exactly what information (e.g., market, product, technology) founders take with them and are able to apply after leaving the parent company.

Spinoff theory predicts that spinoffs are able to use the knowledge gained previously to pursue broad activities that were also undertaken by their parents. For example, Franco and Filson (2006) propose a model in which researchers within a firm imitate their employer's know-how, which implies that the activities of the spinoff will be a subset of those of their parents. Klepper and Sleeper (2005), as well as Klepper and Thompson (2010), also propose that spinoffs will be active in a subset of the broad areas of activity as their parents, but may narrowly depart from the focus of their parents. Similarly, contracting theories also see spinoffs as pursuing activities that originate within the parents (e.g. Anton & Yao, 1995; Cassiman & Ueda, 2006). This leads to the prediction that:

> *Spinoff firms are likely to enter a subset of the broad technological areas where their parents are active at the time the spinoff enters.*

If a spinoff founder's experience has an effect on the firm through knowledge transferred and the previous exposure to the technical and/or market area, we should expect the performance of

spinoffs to be related to the performance of its parent firm. In fact, recent literature finds that spinoff firms with founders from high quality firms in related industries outperform other comparable firms (Klepper, 2007a). While firm performance is related to a number of factors, spinoff founders are influenced by their experience at the parent firm where they were exposed to high performing organizations. As a result, one might expect that spinoffs founded by individuals that come from high performing parents would be more likely to perform particularly well, as stated below:

> *Spinoffs from high performing parents are more likely to out-perform other firms, all else held constant.*

Finally, spinoff theories suggest that better firms spawn more numerous and better spinoffs (Franco & Filson, 2006; Cassiman & Ueda, 2006; Klepper & Thompson, 2006). Since firm founders tend to stay close to their geographic origins (Buenstorf & Klepper, 2009; Dahl & Sorenson, 2009; Figueiredo, Guimaraes, & Woodward, 2002; Michelacci & Silva, 2007; Parwada, 2008), a robust spinoff process could become the basis for an agglomeration of high performing firms clustered around the successful early entrants. This story also describes the development of Silicon Valley.

Similar Results, Different Mechanisms

Both sets of theories describe a process which results in the clustering of a large number of well performing firms operating at the technological frontier with a region (or regions). However, the

mechanisms by which this development occurs are very different, and offer insights with respect to what evidence is expected if either theory was at play in the development of a region.

The following broad predictions relate to agglomeration theory:

- *Firms located within existing clusters are more likely to out-perform firms located outside of clusters following entry, all else held equal.*
- *The rate at which existing firms not producing at the technological frontier subsequently adopt products at the technological frontier at any given moment is greater for producers located in clusters.*

These predictions are based on regional benefits from agglomeration. We would expect to see regional level advantages, both with respect to performance and access to the technological frontier to be evident in the data if agglomeration theory is driving regional development. Given that this theory is agnostic as to whether firm background plays a role in performance and access to the technological frontier, differences with respect to these issues between regions should be entirely explained by regional effects and explanatory variables.

The following broad predictions relate to heritage theory:

- *Existing firms with relevant technical experience have a higher propensity to enter related products, all else held constant.*
- *Existing firms with relevant technical experience out-perform other firms (either new firms or firms without relevant technical experience) following entry, all else held constant.*
- *Spinoff firms are likely to enter a subset of the broad technological areas where their parents are active at the time the spinoff enters.*
- *Spinoffs from high performing parents are more likely to out-perform other firms, all else held constant.*

These predictions are based on effects based on firm backgrounds, with entry, performance, and access to the technological frontier being explained by a firm possessing related experience. Additionally, spinoff firms are likely to benefit from experience gained while the founders of the spinoff were employed elsewhere within the industry. Given the background-specific nature of this theory, it is expected that any apparent regional effects with regard to entry, performance or access to the technological frontier would be explained entirely by the inclusion of explanatory variables representing firm background as long as the backgrounds of the firms are perfectly identified.

The tension between the theories needs to be explored, and the specifics of the industry within which they will be examined may help to do this. The general expectations presented above will be placed within the context of the semiconductor industry in the next chapter in order to formulate specific expectations and hypotheses that can be tested using industry data.

Chapter 3: The Evolution of Semiconductor-based Electronics

The previous chapter discussed how existing theories of firm entry and performance can help to describe the development path of new industries. While these theories provide general predictions, existing research does not make clear the degree to which one might expect each of them to explain the evolution of a particular industry. These theories have been formed from and tested with insights from various industries. But each industry is unique in many aspects, including the location of existing producers, nature of technological change and information, and the relevance of existing capabilities within potential entrants. As a result, the analysis of the semiconductor industry, while considering the general theories, must be informed by its specific characteristics, in terms of firms, technologies and practices. Only then can detailed predictions be formulated, taking into account both generic predictions from existing theories, and the specific context that informs how those predictions can be refined to represent the industry of interest. This chapter explores the development of semiconductor-based electronics in the United States, and using this background, develops and presents detailed hypotheses that will allow for aspects of the theories presented in the previous chapter to be tested.

Many works present the story of semiconductor technology beginning with the invention of the transistor at Bell Labs in 1947. Yet, it is useful to start looking at semiconductors and electronic devices before that to appreciate the succession of knowledge and technology leading up to the transistor. This context sets up the examination of two devices and two corresponding eras in the industry: transistors and integrated circuits. These devices, which both ushered in significant technological progress, presented very different challenges to the industry that will be explored

in detail in this chapter. Ultimately, although building on the same theories described in the previous chapter, differences between transistors and integrated circuits and the state of the industry when each device was invented will result in contrasting expectations and predictions when the industry is examined.

From Cat's Whisker to Transistor

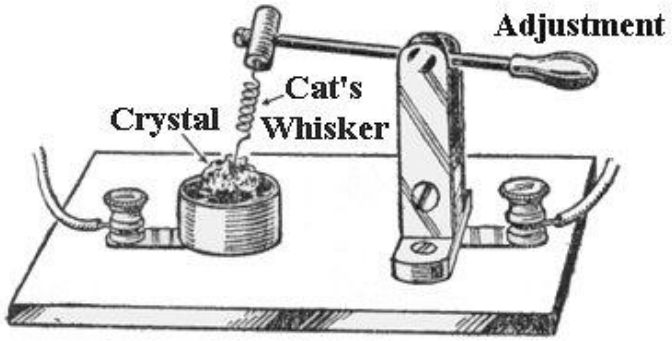
Semiconductors had been used in the early 20th century as rectifiers, known as cat's whisker devices, to receive transmitted signals (through radio transmissions, for example). When transmitting signals using radio waves, two components of the signal are transmitted: an electromagnetic signal (carrier signal) and an electrical signal (input signal). The input signal is modulated to the carrier signal frequency and then transmitted to the receiving device, where the signal is decomposed so that the receiving device can process only the input signal and not the carrier wave which contains no useable information besides the transmission frequency. The technology used to transmit and receive signals has evolved over time to allow for more precise and reliable transmissions,[6] however in the early 1900's, the cat's whisker rectifier was used to perform this decomposition and allow receiving devices to receive the input signal (Braun & Macdonald, 1982 pp. 9-11).

The device consisted of a small semiconductor crystal (galena) that was connected to a metal wire. Different locations on the crystal were conducive to receiving specific signals, requiring

[6] See Aitken (1994) for a discussion regarding interference generated by spark transmitters and the role of interference in motivating spectrum management, and Harrison (1979) for a discussion of tuning innovations that increased the accuracy of transmission equipment.

the operator to determine the placement of the wire by trial and error depending on which signal was desired to be received. When signals were sent through the device and metal wire, the signal was rectified, meaning that electricity could only flow one way (Braun & Macdonald, 1982 p. 11). An illustration of a cat's whisker rectifier is shown below in Figure 1.

Figure 1: Cat's Whisker Rectifier.



Source: http://www.localhistory.scit.wlv.ac.uk/Museum/Engineering/Electronics/history/catswhisker.jpg

While the device became very popular due to its small and inexpensive nature, it had significant reliability issues. These reliability issues were caused by the fact that proper rectification depended on the placement of the metal whisker on the surface of the crystal, which was determined by trial and error.[7] Cat's whisker rectifiers were never produced in large quantities for sale by commercial entities, but rather were produced either in-house by wireless equipment firms or by radio hobbyists for their own equipment (Braun & Macdonald, 1982 p. 12). Over

[7] The cat's whisker design is similar to that used in crystal radio hobby kits. Anyone who has experimented with a crystal radio set can understand how frustrating it would be to mass produce receiving equipment using a cat's whisker diode where the placement of the whisker had to be determined by trial and error.

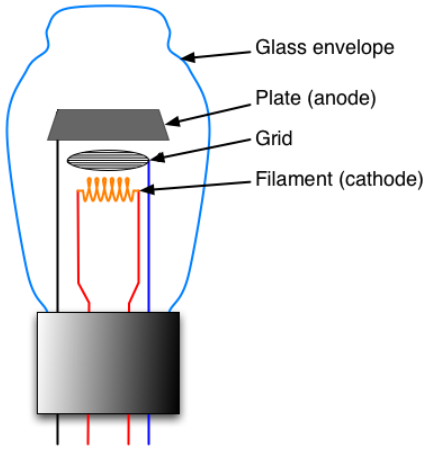
time, the cat's whisker rectifier was replaced by a more robust device that was mass manufactured for sale - the vacuum tube diode.

The vacuum tube diode owes its existence to the light bulb. It took only minor changes in the design and mechanized production of light bulbs to create a vacuum tube diode that would rectify just like the cat's whisker diode and operate at specific "tuned" frequencies. One such change was the addition of a cold electrode (anode) to attract electrons from the hot filament (cathode), as implemented by Fleming in 1904 (Braun & Macdonald, 1982 p. 13). The implementation of a vacuum tube diode resulted in a circuit that was much more reliable than previous cat's whisker diodes as the heating of the filament caused electrons in the filament to gain enough energy to break their bonds and become attracted to the plate (Braun & Macdonald, 1982 p. 13). Its operation had nothing to do with the placement of a piece of wire on a crystal of semiconductor that was open to the air, but rather involved two items that were in a vacuum enclosed by glass. While vacuum tube diodes were more reliable than cat's whisker devices and could handle stronger currents, one disadvantage of the vacuum tube diode was that it required a large amount of power to heat the filament to a state at which it began to emit electrons (Braun & Macdonald, 1982 p. 13).

While the reliability and current processing advantages alone would have most likely been enough for the vacuum tube diode to replace the cat's whisker rectifier, a modification to the vacuum tube diode made vacuum tubes even more attractive to device manufacturers. In 1906, Lee De Forest invented the vacuum tube triode by inserting a grid between the plate and filament of a vacuum diode. The triode – or three-element vacuum tube – allows a small electrical signal

applied to the grid to vary the current flowing between the filament and plate, creating an amplification effect (Braun & Macdonald, 1982 p. 13). An illustration of a vacuum tube triode is shown below in Figure 2.

Figure 2: Vacuum Tube (Triode)



Source: http://upload.wikimedia.org/wikipedia/commons/6/65/Triode_vacuum_tube.png

The vacuum tube triode was able to rectify and amplify electrical signals that were received, as well as oscillating signals that could then be broadcasted. With respect to the received signals, the ability to both rectify and amplify signals with a single device made the device very advantageous to radio receiver producers. In a radio receiver, the rectifier decomposes the received signal so that the device can further process the signal with information (audio sounds). While the raw signal from the rectifier might be able to be played on headphones, the strength of the signal is very weak and not able to be played over speakers without a built-in amplifier. By combining the rectifier and amplifier within a vacuum tube triode, manufacturers of radio receivers were able to use a single device within their receivers to allow for received signals to

be used with a multitude of playback devices (headphones, speakers, etc.). More importantly, this device was reliable (in comparison to cat's whisker rectifiers) and could be mass produced. This innovation facilitated an important growth in the use of radio equipment and other technologies associated with radio (Braun & Macdonald, 1982 p. 13).

Vacuum tubes were used as amplifiers and rectifiers in many products, including telephone repeaters, radio receivers and burglar alarms. However, there was one application of particular importance to the US government where vacuum tubes could not be used – radar detection during World War II (Braun & Macdonald, 1982 p. 14). Radar uses high frequency electromagnetic waves to detect metallic objects. When the electromagnetic wave intercepts a metallic object, the wave is reflected off the object and returns to the transmitting device, where it must then be detected, or rectified. Vacuum tubes are not able to process these high frequency signals because the capacitance of tubes is too high to perform such rapid signal detection (Riordan & Hoddeson, 1997, p. 89).

Frustrated by the fact that neither vacuum tubes or copper-oxide rectifiers were able to detect radar signals, George Southworth, who worked at Bell Labs' field radio laboratory in Holmdel, NJ, decided to give galena-based cat's whisker diodes a try. After searching almost an hour for the proper placement of the cat's whisker, he found that the device was able to detect radar signals (Riordan & Hoddeson, 1997, p. 89). Southworth took his results to a colleague at Bell, Russell Ohl, who had been working on ultra-high frequency/short wave radio topics, and convinced Ohl to pursue a study of crystal detectors to determine what type would be best suited

for radar applications. While Ohl was not entirely surprised by Southworth's results with the cat's whisker, he was surprised by his results with a silicon detector:

> "I tried many kinds of receiving circuits, with peanut-type vacuum tubes and other special vacuum tubes, and none of them worked. So then I got out my old silicon detector and used it. Lo and behold, it was sensitive as the dickens! ... I loaded that up with receiving equipment, and I went all around New York University with it. I was getting strong interference patterns from the elevated line from across the Harlem River on the West Side. There I began to appreciate the power of the crystal detector – this was the silicon detector." (Riordan & Hoddeson, 1997, p. 90)

Following his initial experiments with silicon detectors (i.e., silicon diodes), Ohl was determined to figure out what caused detecting materials to behave in the way that they did. Through his research, he found that the crystal structure of materials in the fourth group of the periodic table were "favorable" (Riordan & Hoddeson, 1997, p. 90) to his purposes. The fourth group of the periodic table contains two particular semiconductors – silicon and germanium – and the group numbers correspond to the number of electrons that each atom in that column has that are available to share with other atoms. Each atom is most stable when it has eight electrons in its outermost shell (i.e., as understood with the Bohr model of the atom), and it is then that the shell is considered "filled". Elements in the fourth column are unique in that they can combine with each other in order to form stable compounds with filled electron shells (Riordan & Hoddeson, 1997, p. 92).

While it was clear that semiconductor detectors could be used for radar in ways that other detectors could not, it was also evident that semiconductor materials were tedious to work with, requiring significant effort to find proper placement of metal connections to the semiconductor for the detector to work properly. Ohl thought the reason for the peculiar behavior of different locations of the semiconductor surface might have to do with impurities (Riordan & Hoddeson, 1997, p. 92). After overcoming some significant difficulties due to silicon's high melting point, which did not allow his team to use traditional melting crucibles, Ohl was able to purify silicon and produce silicon wafers that were much more uniform in terms of their rectification properties. During these activities Ohl's team created a large piece of silicon, known as an ingot, which exhibited unique rectification properties. It appeared that the silicon had some type of barrier which didn't allow for proper rectification. Ohl determined that this "barrier" was something that should be avoided in fabricating silicon detectors, and instructed his team to increase the operating temperature of the furnace in order to produce purer ingots (Riordan & Hoddeson, 1997, p. 93). The barrier problem was ignored by Ohl and his team for almost six months, when it would prove to be a significant finding.

Beginning on February 23, 1940, Ohl's team commenced a line of research that would set the stage for significant advances in semiconductor technology in the coming years. The barrier that Ohl had previously described was actually a separation between regions of the ingot that had distinct impurities. These regions had been formed, inadvertently, as the impurities in the silicon settled slowly during the cooling process, with the lighter impurities rising to the top and the heaver impurities sinking to the bottom of the ingot. In one region the impurity (known as a dopant within the semiconductor community) was phosphorus, which resulted in an excess of

electrons or a negative charge, while in the other region the impurity was boron, resulting in a deficiency of electrons or a positive charge. The negative region was described as n-type and the positive region as p-type by the team, while the natural description of "p-n junction" emerged for the barrier between the two regions. While the junction normally would not allow current to flow through it, when exposed to light, the junction allowed current to flow normally, with the excess electrons (in the n-type area) and the excess holes[8] (in the p-type area) serving as the carriers of current through the device. Inadvertently, the team had discovered the photovoltaic effect of the p-n junction (which allows modern day photovoltaic cells in solar power collectors as well as sensors for light-based switches to operate) (Riordan & Hoddeson, 1997, pp. 94-96).

The discovery of junctions within semiconductors, specifically those in the fourth group of the periodic table, resulted in great interest in the scientific community (Riordan & Hoddeson, 1997, p. 125).[9] Additionally, the realization that these devices could be used for military applications such as radar only added urgency to the need to discover as much as possible in order to transform these discoveries into usable devices for the war that was developing. Working groups emerged between companies, wartime research laboratories and universities to help meet the

[8] A hole is a positive charge carrier that exists when there is a deficiency of electrons (Braun & Macdonald, 1982, p. 22). It can be thought of as the opposite of an electron. Holes are also known as minority carriers, a concept that was not recognized or understood until the work at Bell Labs leading to the discovery of the point contact transistor. For a more complete account of this work and a more detailed discussion of the role of majority (electron) and minority (hole) carriers, please see Nelson (1962) and Pearson and Brattain (1955).

[9] Investigation of semiconductor materials had occurred since Faraday's discovery in 1933 that "the electrical resistance of silver sulphide decreased with increasing temperature while that of other conductors increased" (Braun & Macdonald, 1982, p. 12). However, there was no "systematic, structured and theoretically supported framework of [semiconductor] knowledge" until the 1930's, which prevented the advancement of fundamental understanding of semiconductors until then (Braun & Macdonald, 1982, p. 12).

needs of the nation in semiconductor science. MIT Radiation Laboratory, the University of Pennsylvania and DuPont worked on silicon purification and research (Braun & Macdonald, 1982, p. 58) while General Electric, Sprague[10] and Purdue University worked on advances in germanium (Riordan & Hoddeson, 1997, p. 122). Germanium had emerged, early on, as the material of choice for semiconductor devices due in part to advantages over silicon in refinement and melting (the differences between germanium and silicon will be covered in detail later in this chapter).

Following the war, the belief that science could be harnessed to solve technical and societal problems was widespread. Vannevar Bush had submitted his report, entitled "Science: The Endless Frontier" to President Truman, encouraging the government to take a greater role in promoting government sponsorship of scientific research, especially at universities in the areas of physics, chemistry, engineering and medicine. Also, industrial research organizations like Bell Labs were reorganizing in order to pursue new research in these areas. William Shockley became a leader of the newly formed Solid State Physics program at Bell Labs, and would co-lead a team of world-class scientists and engineers in pioneering work on solid state technology, which also included John Bardeen, Walter Brattain, Gerald Pearson, Robert Gibney and Hilbert Moore (Riordan & Hoddeson, 1997, pp. 115-118). Shockley, who joined Bell in 1936 (Riordan & Hoddeson, 1997, p. 81) had been involved in previous solid state work there and had been influenced by the goal of Marvin Kelly, Director of Research at Bell, to replace electro-mechanical switching equipment used in the Bell system with an all-electronic equivalent.

[10] Sprague manufactured resistors and capacitors for electronic circuits.

Mechanical relay switches and vacuum tubes were the devices that needed to be replaced. Kelly was concerned that ultimately electro-mechanical relays would impede Bell's progress in the future as the system became more complex for two reasons: first, limited switching speed and accuracy of electro-mechanical relays and secondly, high failure rates of these devices in part driven by the reliability issues of the vacuum tube (Braun & Macdonald, 1982, p. 36). The primary weaknesses of the tube resulted from its reliance on a filament to function: the heating of the filament required a great deal of energy, and filaments were fragile and burned out after extended usage, resulting in poor reliability. A semiconductor-based device would not require a filament, and as a result, it would eliminate the known weaknesses, while potentially replacing both the relay and the vacuum tube triode.

While Shockley had previously performed experiments with Brattain to construct semiconductor-based diodes, the group wanted to take its activities one step further to develop a semiconductor switch and amplifier. Early experiments within the group focused on the creation of what became known as a field effect transistor, a device that functioned like a vacuum tube triode when exposed to an electric field on the surface of the device. When these experiments failed, Brattain proposed the explanation that charges at the surface of the semiconductor were blocking the field from having the desired effect, and he generated calculations that indicated that a relatively small amount of charges would be needed at the surface in order to block the field. Bardeen and Brattain began investigating surface charges and by November 1947, after several different approaches, were able to demonstrate the transistor effect by placing two contacts very close to one another on the surface of the semiconductor. The result was the point contact transistor, patented by Bell Labs, demonstrated publicly for the first time in New York

City on December 23, 1947 (Morris, 1990, p. 27). A picture of the first point contact transistor is shown below in Figure 3.

Figure 3: First Point Contact Transistor.

Source: http://homepages.rpi.edu/~schubert/Educational-resources/1947%20First%20point%20contact%20transistor-2.jpg

Transistor Knowledge Access through Government Intervention and Competitive Interests

Due to the potential military implications of the transistor, the device was kept a secret and formally announced only after Bell Labs had briefed representatives from the U.S. Army, Navy and Air Force (Riordan & Hoddeson, 1997, pp. 161-162). However, this sense of secrecy was not something that continued throughout the transistor era. Once Bell secured patents for each invention, access to information about the transistor was relatively easy to get. This ease of access was partially due to government anti-trust mandates placed on Bell Labs and partially due to the competitive interests and licensing practices of makers of transistors, including Bell Labs.

This section will discuss the results of both mechanisms to demonstrate how they worked in tandem to create an atmosphere of openness within an industry that was very competitive.

Government Intervention & Bell Labs

Despite the expectation that a firm that had developed a radically new technology would take extreme measures to protect its intellectual property in an effort to fully appropriate its knowledge, Bell Labs did not take such measures. Rather, the government imposed significant requirements on Bell Labs through two mechanisms: first, through requirements to share information with the government as well as the government's contractors and sub-contractors, and second through a consent decree that was accepted by Bell in 1956 in response to an anti-trust investigation.

Through its research contracts, the government exerted substantial influence on Bell Labs. One example of such an influence was the government's insistence that Bell Labs host transistor technology symposia to share information on transistor uses and features as well as methods of production (Riordan & Hoddeson, 1997, p. 196). Bell hosted three symposia. The first one in 1951 was aimed at military contractors. Following the granting of a patent for the transistor, and in response to a government anti-trust investigation (to be discussed in more depth later), Bell decided to license its transistor technology for a pre-payment of $25,000 in royalties. The second and third symposia in 1952 and 1956, respectively, were held to help transfer transistor technology to Bell's transistor licensees (Moore & Davis, 2004, p. 24; Riordan & Hoddeson, 1997, pp. 195-198; Tilton, 1971, p. 75).

The first symposium was held on September 17, 1951 at the request of the armed services. The event included 300 participants, with 200 being selected by the Army, Navy and Air Force, and the remaining 100 being selected by Bell Labs, mainly among licensees of its transistor patents (Riordan & Hoddeson, 1997, pp. 195-196). At this point, some of Bell's transistor licensees included General Electric (GE), Motorola, the Radio Corporation of America (RCA), Sonotone and Westinghouse (Riordan & Hoddeson, 1997, p. 196). Given the ongoing Korean War, the military was concerned about the secrecy of the device that would surely be used for military applications, while it also desired to inform its staff and contractors about the device that they would likely see in the future. Bell and the military formed a compromise: Bell would hold the symposium to educate the military branches and its suppliers about the transistor, but would not disclose secrets required to manufacture the transistor (Riordan & Hoddeson, 1997, pp. 196-197). As a result, Bell made sure that fabrication technology and methods were not presented or discussed at the symposium, a fact that many attendees lamented (Riordan & Hoddeson, 1997, p. 196). This exclusion of fabrication information would not be the case in later symposia, which were motivated by a desire to disseminate information regarding the state-of-the-art to licensees as opposed to responding to government requirements, and were held after the point at which the military had allayed device secrecy concerns.

In addition to the government's requirement to share transistor knowledge with other military contractors through the transistor symposium, Bell Labs also had substantial restrictions on what types of activities the organization could pursue as a result of the consent decree that was signed in 1956 by AT&T, Bell Labs' parent. The origin of the consent decree was from a lawsuit filed by the Justice Department in 1949 that targeted the monopoly that Bell had over the US

telecommunications system. The lawsuit requested that AT&T divest its manufacturing arm, Western Electric, in an effort to introduce competition into the system through hardware manufacturing. However, the consent decree allowed AT&T to keep Western Electric as its manufacturing arm as long as AT&T agreed to produce transistors only for telephony use within the Bell system and for military purposes, relinquish rights to royalties for any patents issued prior to 1956, license on a liberal basis its inventions to any interested domestic firm, and limit royalties for patents issued after 1956 to reasonable rates (Tilton, 1971 p. 76). While the 1951 and 1952 symposia and other knowledge sharing activities occurred prior to the consent decree being accepted by Bell, it should be noted that the Bell system had been under investigation since 1949, and the organization knew that it had to make knowledge available and perform activities to enhance the greater good so that it would not be portrayed as an evil monopoly (Tilton, 1971, p. 76).

Both of the knowledge sharing activities mentioned above – the 1951 transistor technology symposium and the increased availability of Bell Labs technology licenses – were instigated by government action. One can wonder how things would have turned out without government action. However, evidence is observable that competitive interests alone drove a significant amount of knowledge sharing within the industry.

Competitive Interests & Licensing Practices

Many firms saw the transistor as a direct replacement to, and thus competitor for, vacuum tubes, however it is also clear that some firms realized that they would need to build on knowledge developed by other firms in order to fully develop the transistor (Tilton, 1971 pp. 75-76). Bell

Labs in particular realized that the transistor was a complex device that would require a significant research effort among multiple firms. A Bell VP stated it this way:

> "We realized that if this thing was as big as we thought, we couldn't keep it to ourselves and we couldn't make all the technical contributions. It was to our interest to spread it around." (Tilton, 1971 pp. 75-76)

And act in its interest Bell did, as did as other firms such as RCA. It was the desire of many within the industry that work pursued at a number of firms could be utilized by many within the industry, all working to move forward the technological frontier, and so organizations routinely made sure that their licensing agreements included cross-licensing clauses. These clauses meant that anything that was developed by either firm was included in the license, meaning that any innovations patented by a firm with a Bell license would be available for Bell to use. The use of these innovations through cross-licensing wasn't free – rather, on a periodic basis the two firms would negotiate payments based on the patent portfolios of each firm (Tilton, 1971, pp. 74-76). It does not appear that the cross-licensing practice was really novel as the practice had been occurring within the electronics industry for some time (Burgess, 2008). The novel aspect of information sharing in the transistor era was the extent to which information was shared, which was so aptly declared by a Bell VP:

> "There was nothing new about licensing our patents to anyone who wanted them. But it was a departure for us to tell our licensees everything we knew. (Tilton, 1971 pp. 75-76)

This sharing of "everything we knew" during the transistor era was accomplished through the a policy of essentially open publication at Bell Labs, which resulted in the publication of Shockley's seminal reference text for transistors (1950), among other documents, as well as through two sets of activities: conferences or symposia and site visits.

Conferences or Symposia

Unlike the first symposium, the second symposium included a great deal of information about transistor manufacturing. This meeting was a nine-day symposium, held at Bell Labs April 21-29, 1952 (Riordan & Hoddeson, 1997, p. 197). Attendees included over 100 individuals from the 40 companies that had paid a $25,000 fee as an advance toward future transistor license royalty payments (Tilton, 1971, p. 75). Some of the firms that attended this event included GE, IBM, Globe-Union and Texas Instruments (Riordan & Hoddeson, 1997, p. 197). The proceedings were later published and became the basis for what was known throughout the industry as "Ma Bell's Cookbook" (Riordan & Hoddeson, 1997, p. 197).

Bell's symposia have attracted the most attention in the literature, but were not the only industry symposia related to the transistor. RCA also held symposia as it licensed technology to a large number of firms and used its symposia to share recent progress and new technologies that had been developed by RCA that were available to licensees. Unlike the Bell symposia, where Bell "began throwing theory at them until they were just too fatigued to listen to any more" (Choi, 2007), RCA symposia were designed to be accessible to engineers who were new to semiconductors or currently working with vacuum tubes. The transistor was explained in a way that tube engineers could relate to, various applications of transistors were demonstrated, with

many being implemented in example products such as an all-transistor television, and an assembly line was set up where attendees could participate in the assembly of a germanium junction alloy transistor (Herold, 1983). An RCA employee remembers William Shockley's participation in the assembly line exhibit as one of the highlights of the symposium:

> "One of our greatest sources of pleasure was to have Bill Shockley here and have him put together a transistor by the alloy technique. His comment was that this was the first time he had really made a transistor and that he was very happy to do so." (Heyer & Pinsky, 1975)

While there are no historical accounts of later symposia, RCA's dependence on licensees of its technology for revenue warranted information exchanges in the future.

Following Bell's 1952 symposium, there were a large number of innovations within the transistor product space, with some occurring outside of Bell Labs, but many occurring within Bell Labs and AT&T's manufacturing subsidiary, Western Electric. With this backdrop, Bell chose to hold one final transistor technology symposium in 1956. Again, licensees were invited to Bell's facilities to learn about the state of the art. This time, topics presented included diffusion and oxide masking, as these technologies had been developed by Bell after 1952 (Tilton, 1971 p. 75).

Site Visits

In addition to the symposia, Bell organized a number of other activities to disseminate its knowledge, mostly in the context of visits to the firm. One example is the hours-long workshop that Gordon Teal, a Bell Labs chemist who led a crystal purity project, held explaining how to

build crystal-pulling machines. Attendees were hosted for two days at Western Electric's plant in Allentown, PA, where they saw a transistor production line in operation (Riordan & Hoddeson, 1997, p. 197). Perhaps it is not coincidental that this substantial amount of production knowledge was shared only after William Shockley's patent for the junction transistor was issued in late 1951 (Riordan & Hoddeson, 1997, p. 197).

After the last transistor technology symposium held in 1956, the transfer of knowledge from Bell to other firms continued into the future. Instead of formal meetings with multiple firms, however, Bell began using individual site visits exclusively to consult with both licensees and non-licensees alike. During the late 1950's, Bell used site visits with licensees to transfer knowledge about epitaxial techniques. While detailed statistics are not available over a long period of time, Tilton (1971) indicates that in 1968 alone, approximately 150 representatives of American firms toured Bell facilities. Because of the cross-licensing nature of licensing agreements with Bell, it was in Bell's best interest to consult with licensees and discuss topics at the technological frontier. After all, if a licensee went on to discover something novel and patent it, Bell would have access to the licensees' technology (Tilton, 1971, pp. 74-76).

Hypotheses: Transistor Market Entry and Performance

It is well documented that Bell Labs pioneered the development of the transistor, a device that would eventually replace vacuum tubes in many applications and create new markets for electronics by overcoming previous limitations of vacuum tubes in terms of frequency response and reliability. Additionally, it is clear that Bell and RCA were committed to disseminating information about the transistor – from how to manufacture the device, to potential applications

– information that could be used by firms interested in producing transistors. However, it is still not apparent how the actual development of the transistor industry fit with existing theories regarding firm entry and performance. What role did the concentration of information within Bell Labs and other leading firms play in the development of the industry, and which theories best describe predictions that match the process that was observed? This section examines the theories discussed in Chapter 2, in light of the observed characteristics and evolution of the industry, to generate hypotheses using these theories in conjunction with what is known about the development of the transistor.

The demonstration of the transistor by Bell Labs in 1947 (Morris, 1990, p. 27) provided a great stimulus to the electronics industry as a whole. As previously discussed, the vacuum tube had substantial limitations with regard to frequency response and reliability. These limitations ultimately motivated Bell to research other devices that could be used in lieu of vacuum tubes for high frequency applications and other uses that required a higher degree of reliability. Transistors addressed both shortcomings of the vacuum tube, making existing vacuum tube producers fearful that their devices would quickly be replaced by the new semiconductor device (Braun & Macdonald, 1982, p. 47). This fear was not unfounded, as estimates indicated that by 1961 over half of tube usage dedicated to television sets would be replaced with transistors (Harrigan & Porter, 1979, p. 6). Tube manufacturing firms were driven to consider producing transistors in the face of such statistics in order to avoid obsolescence.

Transistors offered a clear path forward for tube producers in replacing devices that had been prone to significant reliability issues. Simultaneously, awareness about the importance of the

device was conveyed through industry predictions and Bell's symposia. Consequently, rectifier and diode producers realized that they too might be able to produce transistors by utilizing their existing semiconductor production knowledge (Braun & Macdonald, 1982, p. 47). Such related technical experience has been shown to positively affect a firm's propensity to enter other industries (Klepper & Simons, 2000; Lane, 1989; Scott-Morton, 1999), and it seems that this related experience may have been of particular use given the availability of technological and manufacturing knowledge from Bell Labs and other pioneering firms through symposia and licensing. Given detailed firm background information, the influence of previous experience can be tested through the following hypothesis:

Hypothesis 1: Among existing firms at the time that transistors were invented, firms with related electronics experience, especially tube manufacturers who faced possible obsolescence from the introduction of the transistor, were more likely to begin producing transistors, all else held constant.

The manufacturing processes involved in producing vacuum tubes are very different than those used to produce transistors. However, there are similarities in both the market and aspects of manufacturing between tubes and transistors that may have afforded advantages to existing tube producers that chose to produce transistors. Given these similarities, it seems reasonable to consider the possibility that a firm's prior related experience in tubes may have been particularly salient to its success in producing transistors.

In the most basic sense, transistors could be used as a direct replacement for vacuum tubes with respect to amplification and switching, as indicated by the earlier statistic regarding transistor usage in televisions. This is not to say that transistors could be swapped in for tubes without any circuit adjustments. Rather, circuit designs needed to be biased appropriately to ensure the

correct operation of transistors. However, the transistor could directly replace the vacuum tube in terms of the operation that it was performing. As a result, there was significant market similarity between vacuum tubes and transistors, which allowed many existing vacuum tube firms to utilize their complementary assets (sales channels (Harrigan & Porter, 1979, p. 6), marketing, etc.) that they had previously used for vacuum tubes when producing transistors. Presumably, these existing manufacturers were able to use these complementary assets to their advantage in producing transistors. A number of authors have shown that incumbent firms have been undermined by technological change that disrupts existing industries allowing new firms to assume leadership positions (Christensen, 1997; A. C. Cooper & Schendel, 1976; R. Henderson & Clark, 1990; Tushman & Anderson, 1986). However, it appears that this occurrence is far from universal. For example, in a study examining the typesetting industry, Mary Tripsas showed that in the instances when technological change did not devalue complementary assets of incumbent firms, these firms were able to retain a vast majority of market share following the technological change. In the instances where these specialized complementary assets were devalued, however, incumbent firms no longer enjoyed any advantages over new entrants and were overtaken by them in terms of market share (Tripsas, 1997, p. 133). Given the applicability of complementary assets possessed by vacuum tube producers to transistor sales, marketing and distribution, it seems reasonable to expect that existing vacuum tube firms would have an advantage with respect to spinoff firms and other new firms in terms of performance in the transistor market.

Many of the specifics of manufacturing transistors were not similar to those involved in manufacturing tubes, but there are aspects that are similar between the two products. First,

production of both products required clean and dustless environments to avoid contamination which would cause the devices to malfunction (Matthews, 2003, p. 134; Harrigan & Porter, 1979, p. 5). Additionally, many of the manufacturing steps required high precision (Harrigan & Porter, 1979, p. 4; Lécuyer, 2006, p. 221). The creation of a vacuum takes very different skills than the diffusion of dopants (or impurities) into semiconductor; however, both tasks require high precision in their performance, and this focus on precision can be transferred to the other tasks required to produce new products. Finally, many persons involved in manufacturing tubes admit that although the production was fairly standardized, "there was a certain element of 'black magic' involved" in their production (Harrigan & Porter, 1979, p. 5). Likewise, this aspect of 'black magic' would be found in the manufacture of transistors, and firms that were experienced in such aggravating manufacturing processes may have been more successful in adopting transistor production.

Taking into account the similarities in markets and aspects of manufacturing, it seems reasonable to expect that firms with prior tube production experience would be more likely to perform well following their entry, since tubes had some similarities to transistors. Such related technical experience has been shown in many cases to positively affect a firm's propensity to enter other industries, and to positively affect performance following entry into those industries (Klepper & Simons, 2000; Lane, 1989). However, it is not clear that the experience gained by tube producers in particular was enough to give these firms any distinct advantage compared to other entrants. This is due in part to the relatively open nature of transistor information from Bell Labs, which may have given other electronics and diversifying firms equal footing with respect to entry into transistor production.

Given the support in previous literature and the technology and market context previously discussed, the notion of related experience playing a role in performance following entry may thus be something that is expected to be observed in the data. This influence can be tested with the following hypothesis:

Hypothesis 2: Among transistor producers, firms with related electronics experience, specifically tube manufacturers, are more likely to perform better than other firms, all else held constant.

Clearly experience is important, but the incentive for tube producers to begin manufacturing transistors should not be overlooked when considering performance of diversifying firms into transistors. With estimates of more than half of tube applications moving to transistor technology in a relatively short time period, it seems that many tube producers were almost forced to produce transistors in order to survive, and to do so quickly. Early producers of transistors were able to enjoy first-mover advantages, as described by Lieberman and Montgomery (1988), specifically technological leadership advantages. These first-mover advantages allowed firms to gain knowledge and experience regarding transistor manufacturing, giving these firms the ability to decrease the cost of production. Reduced production costs could then be re-invested in further research and development, which might result in future products that would allow the firm to increase their market share, or passed on through price reductions that may increase the market share of the firms (Lieberman & Montgomery, 1988, p. 42). Therefore, familiarity with tube technology and markets may have been important resources for firms considering entry into transistors. However, the timing of entry may have played a more important role in which firms would amass large market shares in what was a quickly developing product space given the differences in manufacturing between tubes and transistors.

The possible existence of an early-mover advantage could thus be examined using the same industry-wide firm data to determine which of the two mechanisms, experience and timing, was influential in the transistor era. The following hypothesis would be tested:

Hypothesis 3: Among transistor producers, firms that begin producing transistors earlier are more likely to perform better than other firms, all else held constant.

Immediately following the introduction of transistors, Bell was the only organization that possessed the cutting-edge knowledge required to produce these devices, since it had pioneered the development efforts. This expertise and knowledge was necessary for any firm that was looking to produce transistors. However, as previously mentioned, Bell was very forthcoming regarding this information, holding symposia and inviting firms to visit in order to learn how to design and manufacture transistors, which made this knowledge widely accessible relatively quickly. Bell disclosed information to a number of firms, most of which had previously produced electronics and were located within existing electronics clusters: Boston, Los Angeles and New York City. Boston had developed into a cluster of electronics producers that were active in aerospace and military applications during World War II (Saxenian, 1994, pp. 12-17), while Los Angeles and New York City were existing clusters for radio production (Klepper, 2007b, p. 2).

As the new technology grew, informal collaborations driven by discussions and interactions between employees of firms pursuing similar products and technologies, or the mobility of employees across firms are predicted to become important vehicles to access valuable knowledge about technology and markets. As explained in the previous chapter, clustering helps such mechanisms in a region. The existing firms, mostly located within electronics clusters,

collaborated on the transistor development work using germanium and silicon; both Bell and RCA held symposia to disseminate information regarding advancements in transistor technology. All of these efforts increased the amount of knowledge available to firms that were interested in developing transistors. Given the presence within the clusters of knowledge sharing activities, the presence of Bell licensees and the knowledge exchanges associated with agglomerated regions, it seems reasonable to expect that firms located within the three electronics clusters should be at an advantage with respect to firms located outside of the clusters. There would be access to applicable transistor knowledge through spillovers between firms that licensed transistor technology from Bell in each of these major regions. Using performance data for transistor producers, the presence (and effects of) regional agglomeration economies within electronics clusters can be examined by testing the following hypothesis:

Hypothesis 4: Among transistor producers, firms located in Boston, Los Angeles and New York are more likely to perform better than firms located elsewhere, all else held constant.

Despite this prediction by agglomeration theory that regional effects existed within agglomerated regions, it is important to consider that the openness of Bell Labs may have reduced any distinct advantage that firms located in New York City and other electronics centers may have otherwise enjoyed through agglomeration economies. Through its licensing, symposia and other technical visits, as well as the firm's emphasis on disseminating information on the cutting-edge of transistor technology, Bell Labs provided firms located throughout the United States with access to information that would have otherwise been more geographically limited to the New York area. Therefore, while agglomeration economies may have still been present within the electronics clusters, the technological information that Bell was providing to firms located

throughout the United States may have decreased the value of being located within such a cluster.

The Role of Spinoffs in Transistor Production

As described in the previous chapter, spinoff firms are firms formed by individuals previously employed within the industry. Heritage theory posits that spinoff firms will be influenced by the employment history of the firm founder(s), both in terms of production as well as performance. Additionally, heritage theory predicts that agglomerations might occur around early, highly-successful entrants as the spinoff process triggers individuals from these leading firms to leave and form spinoffs nearby, most of which are also expected to perform well (Klepper, 2009b, p. 163). Theory suggests that spinoff firms will enjoy advantages over other new entrants, and in some cases existing firms, in several areas, including access to technological knowledge and market knowledge. This section will examine the specifics of the transistor era of semiconductor electronics in order to form hypotheses using both heritage theory and the context of the industry.

Technology Knowledge

As documented, knowledge regarding transistor technology was readily available from Bell Labs. The organization, under extreme anti-trust pressure, took every action possible to disseminate "everything they knew" (Tilton, 1971 pp. 75-76) about transistors, transistor technology, and transistor manufacturing processes. As a result, employees of Bell looking to found their own firms might not have had any decided advantage in terms of access to transistor technology knowledge when compared to their peers employed elsewhere. Contrary to most

other cases where spinoffs have been found to emerge, critical information was accessible through interactions with Bell researchers through symposia and visits, as well as through Bell licenses. This is in direct contrast to industries such as lasers (Klepper & Sleeper, 2005) and hard drives (Franco & Filson, 2006). In these cases, the founders of spinoff firms utilized critical and unique information gained in their previous employment to begin production. As such, their processes were based on the knowledge acquired from the parent firm, information that was virtually inaccessible by other means. It remains to be seen, however, whether knowledge regarding transistor design and manufacturing gained from previous employment at Bell was less relevant in the quest to produce cutting-edge transistors since Bell Labs served as a ready, open source of transistor technology knowledge for the industry.

Market Knowledge

With some re-design of circuits, transistors could be used as a direct replacement for the functions being performed by vacuum tubes (amplification and switching), as indicated by the earlier statistic regarding potential transistor usage in televisions. As a result, there was significant market similarity between vacuum tubes and transistors, which allowed many existing vacuum tube firms to utilize their complementary assets (sales channels, marketing, etc.) that they had previously used for vacuum tubes when producing transistors. As discussed previously, literature has shown that while technological change can disrupt existing industries causing diversifying firms to falter (Christensen, 1997; A. C. Cooper & Schendel, 1976; R. Henderson & Clark, 1990; Tushman & Anderson, 1986), such firms are able to retain significant market share following technological change if the complementary assets they possess are not de-valued by the technological change (Tripsas, 1997, p. 133). Given the applicability of complementary

assets possessed by vacuum tube producers to transistor sales, marketing and distribution, it seems reasonable to expect that existing vacuum tube firms would have an advantage with respect to spinoff firms and other new firms in terms of performance in the transistor market.

Within the literature, references to transistor spinoff firms are virtually non-existent when compared with discussions of later integrated circuit spinoffs. There are several notable spinoff firms during the transistor era, specifically Shockley Semiconductor and Transitron from Bell Labs. However, diversifying tube producers achieved and retained significant market share for transistor products (Tilton, 1971, p. 66). When examining transistor innovations, it is notable that until the advent of the mesa and planar transistors, which revolutionized the design and manufacturing of the transistor, diversifying firms such as GE and RCA created the vast majority of innovations. But why is it that spinoff firms did not play a large role in transistor innovations and failed to perform well given the predictions of heritage theory? It appears that the lack of spinoff advantage through tacit or secret technical and market knowledge may explain the dominance of diversifying firms.

In examining the performance of spinoff firms, it appears that spinoffs in the transistor era might not have enjoyed many advantages as compared to other firms. As a result, it might be expected that the relative number of spinoff firms in transistors would be small and that the performance of these spinoff firms would not be particularly notable as compared to other firms. The performance prediction can be examined by testing the following hypothesis:

Hypothesis 6: Among transistor producers, spinoff firms should perform in a similar to fashion to other comparable firms, all else held constant.

From the Point Contact Transistor to Integrated Circuits

While Bell Labs led the initial transistor development, the introduction of the device set off a flurry of activity among electronics producers who were trying to catch up with, and possibly overtake, Bell in terms of semiconductor technology. Through this competition, a new type of device would be developed called the integrated circuit. This new device was notable for a number of reasons. First, the improvements in reliability would allow electronics to be used in increasingly critical applications, such as the Apollo guidance computer.[11] Additionally, the complexity that could be achieved by using integrated circuits created new applications and markets for electronics, exposing many consumer electronics makers and ultimately end-users to semiconductors for the first time. Through these two advancements, the level of opportunity involved with the advent and commercialization of the integrated circuit was unprecedented, attracting a large amount of commercial interest. Finally, the integrated circuit was especially notable because it made possible the growth of the region now called "Silicon Valley". The region, previously known for its apricot groves, was transformed by the advent of the integrated circuit into the "capital" of the electronics industry. The development of the integrated circuit, which set up the locational shift of the industry, is best understood by examining the intermediate semiconductor devices that were developed leading up to the integrated circuit. This section details transistor developments after the introduction of the point contact transistor, and

[11] See Mindell (2008, pp. 130-133) for a discussion on the reliability requirements of the Apollo guidance computer and what lessons were learned by IC producers to increase device reliability.

showcases the role of firms other than Bell in the development of these devices in the lead up to the invention of the integrated circuit.[12]

While the point contact transistor had been a major step forward for solid state electronics, the device invented by Bell Labs had several weaknesses. Specifically, the placement of the contacts on the surface of the semiconductor material was tenuous at best. Indicative of the reliability issues with the point contact transistor was the fact that the first device could only be made to work "by Water Brattain (co-inventor) and only, he wrote later, 'if I wigged it just right'" (Reid, 2001, p. 55). Additionally, the range of frequencies at which the transistor would operate was small compared to vacuum tubes (Morris, 1990, pp. 29, 42). For the transistor to succeed as a device to replace vacuum tubes, further development was needed. In addition to Bell Labs, General Electric and RCA were some of the early firms invested heavily in transistor research and development, which resulted in the invention of other types of transistors. However, the development path for General Electric and RCA was a bit uncertain – not only because of uncertainties in the technology, but also due to uncertainties regarding Bell's patenting and licensing strategy. Bell would eventually adopt a very liberal licensing policy, but this was not clear until the early 1950's. As a result, firms such as RCA initially attempted to understand what Bell had been doing while "inventing around" them (Burgess, 2008). Although

[12] While the goal of this section is to present a comprehensive discussion about transistor development following the point contact transistor, it is by no means exhaustive. Several transistor variants (surface barrier and drift transistors, for example) are excluded from the discussion, though the discussion of all major transistor families that were introduced up to and including the planar transistor has been attempted. MOSFETs have been excluded from this discussion as they were developed significantly later than most of the technological developments of the transistor era. The analyses of transistor entry, performance, etc., do not include the time period when MOSFETs were invented.

this doesn't seem to have had any effect on the invention of grown junction transistors, it explains why we see firms besides Bell at the forefront of developing other transistor families in the early years of the industry.

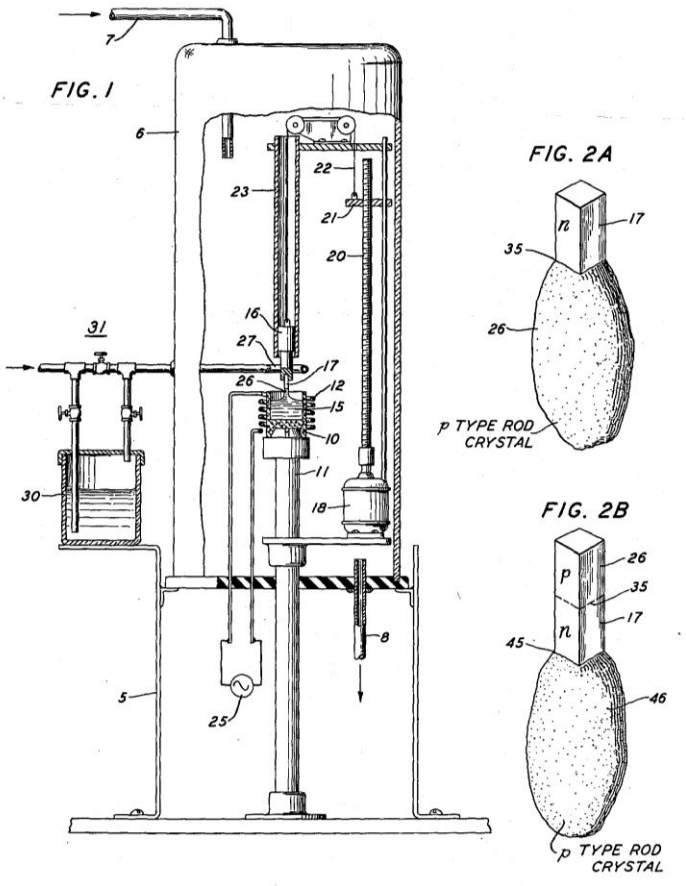
Germanium Grown Junction Transistor

The invention of the grown junction transistor is rooted in the work of two individuals. William Shockley is credited with describing the design and function of the junction transistor as early as 1948 (Morris, 1990, p. 29). However, it was Gordon Teal's interest in high quality semiconductor materials that allowed the device to be developed and produced. Teal felt that the multi-crystalline nature of many semiconductor materials contributed to the impurities problem, and that the problem could be solved, in part, by growing large single crystals for use in fabrication. Others within Bell Labs, including William Shockley, did not see the problem the same way, and discouraged Teal from investigating the topic by declining requests for funding and time to work on such problems (Riordan & Hoddeson, 1997, pp. 173-174).

Fortunately for Teal, an opportunity arose to produce a small scale germanium crystal in order to test out his ideas regarding crystal growing. On his way to catch a bus to another Bell facility, Teal ran into John Little, a Bell Labs mechanical engineer working in device development, who was looking for a thin germanium rod to be used to produce a filamentary transistor. Sensing this would be the perfect opportunity to demonstrate his ideas regarding single crystal growing, Teal offered that he could make a thin rod from a germanium melt, and indicated that it would also be a single crystal. Using a crystal growing method invented in 1916 by Jan Czochralski, a Polish scientist, Teal and Little designed a growing mechanism to produce the germanium rod

that Little had in mind (Riordan & Hoddeson, 1997, p. 174). The apparatus used a seed crystal, which when placed in contact with molten germanium would attract the germanium. The seed crystal would then be slowly retracted from the molten liquid, and as the liquid was pulled upward it would cool into a solid, single crystal (Riordan & Hoddeson, 1997, p. 174). An illustration of a later version of the crystal growing apparatus is shown below in Figure 4.

Figure 4: Teal's Crystal Growing Apparatus.



Source: http://patft.uspto.gov/netacgi/nph-Parser?Sect2=PTO1&Sect2=HITOFF&p=1&u=%2Fnetahtml%2FPTO%2Fsearch-bool.html&r=1&f=G&l=50&d=PALL&RefSrch=yes&Query=PN%2F2631356

Within a few days, Little had the germanium rod that he needed, and Teal, having shown that the proposed method could produce single crystal germanium that could be used to construct transistors, had the proof of concept that he needed to appeal again to his supervisors. But his bosses and Shockley again turned him down. Finally, after one last request to Jack Morton, the head of the device development team, he got some funding to build (but not operate) the crystal pulling apparatus (Riordan & Hoddeson, 1997, p. 174).

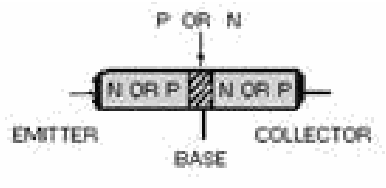
Over the next year, Little and Teal built a full-size apparatus measuring 7 feet tall and performed further crystal growing experiments. Since the activity was not funded in terms of operation, Teal worked double shifts for most of 1949, the only way he could focus on his true passion (Riordan & Hoddeson, 1997, p. 179). For that purpose, the apparatus was placed on wheels so that it could be wheeled into the lab each night and back into storage early the next morning (Riordan & Hoddeson, 1997, pp. 178-179). Teal recalled the process that was followed each day in order to return the lab to its proper setup:

> "This meant that frequently around 2 or 3 o'clock in the morning I had to disconnect from the wall approximately 30 foot hydrogen, nitrogen and water-cooling lines leading to the puller as well as high-power electric lines to the high-frequency heater." (Riordan & Hoddeson, 1997, p. 179)

Teal's efforts were rewarded when the wafers created from the single crystal germanium had substantially better performance, and Shockley admitted that he had been wrong about dismissing Teal's idea (Riordan & Hoddeson, 1997, p. 179). The growing process was modified to produce a transistor by first doping the molten germanium with a p-type impurity, pulling part

of the ingot, then dropping in n-type impurities followed by p-type impurities as the crystal was continuously pulled in order to form a PNP "sandwich". The crystal was later cut into bars and leads were attached to the three regions, resulting in the grown junction transistor (Morris, 1990 p. 30). The grown junction transistor was first introduced in 1951, and by 1952 Bell was producing 100 NPN transistors each month (Morris, 1990 p. 31). An illustration of a grown junction transistor is shown below in Figure 5.

Figure 5: Grown Junction Transistor.



Source: http://www.rfcafe.com/references/electrical/NEETS%20Modules/images/07241img1A.gif

The grown junction transistor was significantly less noisy than the point contact transistor. However, the increased width of the point junction's base as compared to the point contact transistor resulted in a lower frequency response, making this performance aspect of the point junction transistor inferior to the point contact transistor (Morris, 1990 p. 32). It was very difficult to control the base width with the growing apparatus, and the minimum base width was limited by the fact that a metal lead had to be mounted onto the base (Morris, 1990 p. 30). Despite its limitations, the grown junction transistor was accepted as a far superior transistor than the point contact (Morris, 1990 p. 31). The impact of this alternative design is reflected in this description in the response from the military:

> "Due to the difficulty of quantity production of transistors with uniform characteristics, immediate application to military equipments [sic] was impractical up to the spring of 1951. Evolution of the junction transistor in 1951 plus quality production ability with uniform characteristics for both these new type of low noise units as well as somewhat older point-contact types, changed the somewhat restricted approach by military services to an active and expanding program." (Morris, 1990 p. 31)

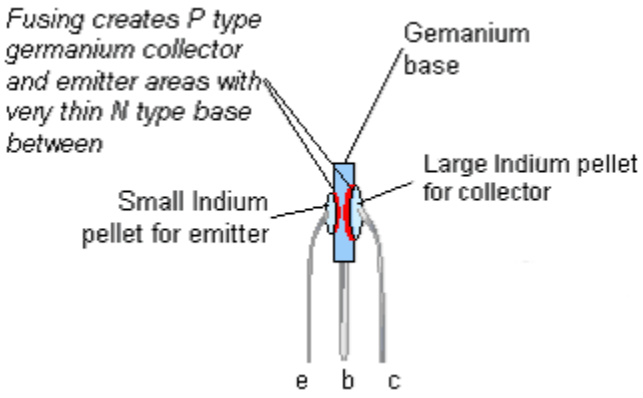
Germanium Alloy Junction Transistor

While the germanium grown junction transistor was more robust than the point contact transistor, its relatively large base width resulted in a much lower frequency range than the point contact transistor. Many tube and semiconductor firms were eager to develop transistor devices and advance the technological frontier, but there was significant uncertainty with respect to the intellectual property rights that would be asserted by Bell Labs. As a result, many firms "invented around the Bell Labs' patents" (Burgess, 2008). This approach of avoiding the initial Bell innovations led GE to develop the alloy junction process in 1952 as a way to reduce the width of the base and, as a result, increase the frequency range of the transistor (Morris, 1990 p. 32). Alloy junction transistors are constructed by attaching two indium pellets onto a germanium base and re-crystallizing the germanium using an alloy furnace at approximately 600 degrees Celsius (Morris, 1990 p. 32), thereby diffusing minority carriers into the base.

While the p-n junctions are exposed in this design, which allows for potential contamination and reliability issues, the reduced base width allowed for a greater frequency range with respect to grown junction transistors (Morris, 1990, p. 34). Additionally, the alloy junction design lent

itself to large scale-production, more so than grown junction transistors, since alloy junction transistors were not formed during the pulling process, resulting in a lower cost compared to grown junction transistors (Morris, 1990, p. 33). An illustration of an alloy junction transistor is shown below in Figure 6.

Figure 6: Alloy Junction Transistor.



Source: http://www.learnabout-electronics.org/images/image312.gif

General Electric had been producing semiconductor diodes since 1948 and had developed alloy junction rectifiers in 1949. Given GE's technological capabilities with respect to the alloy process as well as growing germanium for use in alloy junction diodes, the alloy junction transistor was a natural extension to its previous development work (Burgess, 2011a). Alloy transistor development was started in February 1951 by John Saby, and the first transistor was created on March 11, 1951. This rapid development timeline stemmed from Saby's not having to develop new technologies. Rather, he was able to take advantage of existing GE technologies to develop a new device (Burgess, 2011a).

While GE is commonly credited with the invention of the junction alloy transistor, RCA actually filed the patent one day prior to GE, and was originally granted the patent (Burgess, 2008). Examining the development of alloy junction transistors at RCA, however, speaks to the difficulties that some firms without previous semiconductor experience had in the early semiconductor era. Rather than drawing on previous production experience like GE, RCA had to learn from published documents and conferences, and there was very little published about alloy junctions in 1951 when RCA began working on them (Burgess, 2008). At the time, RCA's transistor development production had been limited to point contact transistors (Burgess, 2008). Additionally, the publications that were available were vague in terms of manufacturing specifics, with the seminal publication from GE's Hall and Dunlap indicating the following about the process:

> "the non linear impurity distribution which is required may be obtained by thermal diffusion of donor and acceptor impurities into opposite sides of a wafer of semiconductor" (Hall & Dunlap, 1950)

Given that even the type of impurities were not specified in the document, RCA researchers had to guess which impurities to use and how to go about the diffusion process (Burgess, 2008). Following six month of failed efforts, RCA began working on diffusion experiments instead. Only after seeing a GE presentation of alloy junction transistors did RCA revive its efforts to produce alloy junction transistors and dropped all development work on point contact transistors. RCA again began working on alloy junctions in September 1951 and had produced 40 junction devices within the month. By December 1951, RCA was confident that it could produce

transistors using its diffusion process, and by April 1952 it had produced pilot runs of PNP alloy junction transistors (Burgess, 2008). While RCA was able to learn the requisite knowledge and overcome the challenges involved with producing alloy junction transistors, it ended up introducing these transistors more than one year after GE.

Silicon Alloy Junction Transistor

Following the development of germanium alloy junction transistors, several firms within the industry began pursuing silicon alloy junction transistors. This pursuit was aimed at overcoming the two key disadvantages of germanium. First, germanium transistors were not able to operate properly at high temperatures due to the lower melting point of germanium as compared to other semiconducting materials, such as silicon (Riordan & Hoddeson, 1997, p. 221). Additionally, the leakage current in germanium was a source of problems for early users of germanium transistors. In short, leakage current allows a non-negligible amount of current to move through the transistor even when it is in the "off" state, meaning that for switching applications, the switch is not truly off (Riordan & Hoddeson, 1997, p. 221). One can only imagine how this could cause issues in computing and other defense-related applications, for example. Silicon, however, did not exhibit this leakage current problem and thus was a preferred material for transistor development (Riordan & Hoddeson, 1997, p. 221). It does not seem overly surprising that firms looking to possibly switch from germanium to silicon pursued silicon transistors using the same methods that had recently been used to move the state of the art forward with regard to germanium transistors – alloy junction transistors.

Silicon alloy junction transistors were made using essentially the same process as germanium alloy junction transistors, and they had many of the same reliability issues due to the design of alloy junction transistors (see previous section on germanium alloy junction transistors for details). While silicon grown junction transistors presented significant hurdles in fabrication, which many firms would find insurmountable based on their knowledge and skills, silicon alloy junction transistors also presented substantial challenges. Specifically, the expansion that occurred when the transistor was heated occurred at different rates between the silicon and the alloying material, which presented challenges related to the robustness of the device (Goldstein, 1991). While firms had the option of pursuing silicon transistors using either grown junction or alloy junction techniques, most firms decided to pursue alloy junction transistors (Goldstein, 1991).

Though many difficulties existed in fabricating silicon transistors, industry interest in the material was clearly growing. During an Institute of Radio Engineers[13] (IRE) conference in Minneapolis in June 1954, an entire section of sessions specifically focused on silicon was held – the first time that this had occurred (Burgess, 2008). RCA took this opportunity to unveil its silicon alloy junction transistor, the first such transistor that had been revealed to the industry. However, although RCA was the first firm to present a functioning alloy junction transistor, it was not the only firm that had been working on silicon alloy junction transistors. At the IRE

[13] The IRE is one of the two organizations that were merged to form the Institute of Electrical and Electronics Engineers, known commonly as IEEE.

conference, both Hughes Aircraft's research unit and Bell also presented on silicon alloy junction development (Burgess, 2008).

While RCA was the first company to develop a silicon junction alloy transistor, the firm chose to abandon the project, and never commercialized it (Burgess, 2008). This decision seems to have resulted from a misguided understanding of how silicon would be received. In his 1990 piece about the "human side" of electronics development, Jack Saddler, an RCA engineer, noted that a marketer had performed a market analysis for silicon transistors. The study showed "that RCA should avoid the material because it would only be useful for military purposes… [and that] the armed forces ultimately would be only a small portion of the market" (Saddler, 1990). As would be evident in later years, this marketer was very wrong about how things would turn out. The study was very influential in leading RCA down a path where it would continue to pursue only germanium transistor production.

RCA's decision to abandon silicon alloy junction transistors was evident in the paper that was presented at the IRE conference. Unlike previous publications that were very guarded with details about fabrication, the RCA paper on silicon alloy junction transistors provided "comprehensive process information, indicating that RCA was not seeking to protect its commercial position around silicon alloy development" (Burgess, 2008). Ultimately, the decision to abandon silicon was an error on the part of RCA, as silicon would become the primary material used in transistor – and integrated circuit – production. While the individual that produced the marketing study "soon found himself out of the semiconductor industry and into steel fabrication" (Saddler, 1990), RCA found itself at a disadvantage for following the

guidance of this individual. Silicon would become the material of choice within the industry, and the market soon shifted to a new type of transistor - silicon grown junction.

Silicon Grown Junction Transistor

For several years, it had become clear that the operating temperature limitation of germanium devices at near 90°F would be an issue for military applications such as rocketry. Silicon, with a maximum operating temperature of near 150°F, was seen as a more desirable material (Goldstein, 1991, p. 31; Seidenberg, 1997). While the increased operating temperature range was a key concern for many producers, silicon had its disadvantages with respect to germanium that had kept many firms from successfully moving forward with a silicon program. Given silicon's higher melting point, traditional crucibles and melting equipment could not be used with silicon (Moore, 1998, p. 35). Additionally, silicon was much more susceptible to contamination compared to germanium (Riordan & Hoddeson, 1997, p. 208), and carrier mobility was lower in silicon than germanium, resulting in slower operation (or a lower frequency response) for comparable devices (Chen, 2004, pp. 12-13).

While these challenges in working with silicon may have deterred many firms, several firms working with silicon were in pursuit of creating the next type of transistor. It should be noted that many of these firms continued to develop and produce germanium transistors as well as

pursuing silicon transistors.[14] RCA and Hughes were both working on manufacturing silicon alloy junction transistors, while Raytheon was working on producing silicon grown junction transistors, and Bell was working on both types of silicon transistors as well as silicon diffused transistors (Burgess, 2011b, p. 23). Even so, there were very few firms that had begun work on a silicon transistor program, due to a lack of knowledge about the material as well as the many challenges associated with it. Those that had begun working with silicon had not announced any significant progress by the end of 1952. However, the actions of one man from Bell Labs – Gordon Teal – changed the pace of developments when he left Bell Labs and announced the first commercial production of silicon transistors shortly after his departure.

Following his successful program to grow single crystal ingots of germanium, Teal had moved on to a new problem – the challenge of growing and doping silicon crystals. Teal would later report that his motivation regarding silicon transistors at Texas Instruments (TI) was based on the potential military uses of such devices, though his motivation for moving from germanium to silicon while at Bell is not clear (Riordan & Hoddeson, 1997, pp. 207-208). By 1951, he had succeeded in growing single crystal silicon with p-n junctions, had published the results in early 1952 (Riordan & Hoddeson, 1997, p. 207), and was almost ready to grow a full-fledged transistor using this method by late 1952 (Goldstein, 1991, p. 30). Teal, however, was growing restless at Bell, with a desire to take on more responsibility within the organization. His restlessness, his wife's homesickness for Texas, and a classified advertisement for a research

[14] Texas Instruments, for example, the leading firm for silicon grown junction transistors, had an active germanium transistor program, and continued to use germanium transistors in projects such as the pioneering Regency TR-1 radio due to cost concerns (Burgess, 2011b).

director position at TI culminated in a move that had significant repercussions for the location of future research and development of silicon transistors (Riordan & Hoddeson, 1997, p. 206).

Teal's initial discussion with TI's executive vice president and director, Mark Shepard, about the Research Director position at TI was not his first interaction with the firm. Shepard and other TI representatives were present at Teal's 1952 transistor technology symposium presentation at Bell Labs, and Teal had traveled to TI following the symposium to teach the group about the pulling process and to explain how the pulling devices should be built (Burgess, 2011b, p. 34-35). Teal later reflected on the fact that his instructions were instrumental in allowing the TI group to build the required pulling machinery:

> "It wasn't done absolutely like I told them, but I am certain I was partly responsible for it turning out successfully… They needed the background knowledge that I already had because I had already designed the damned machine… I think we had four machines working before I left there and we'd done a lot of experimenting in building those machines in different ways." (Burgess, 2011b, p. 34-35)

While the TI group hadn't heeded all of his instructions in building the pulling machinery, the group had early success with the equipment. Using crystal seeds provided by Bell, TI was able to successfully pull germanium crystals and produce grown junction germanium transistors in 1952, prior to Teal's arrival (Burgess, 2011b, p. 4). Silicon grown junction transistors, on the other hand, had not yet been pursued by the organization.

Bringing with him all of his experience and knowledge regarding crystal growing and the grown junction production process, Teal joined Texas Instruments on December 31, 1952 (Goldstein, 1991, p. 27). Immediately, Teal set out to produce the first grown junction transistors made of silicon (Riordan & Hoddeson, 1997, pp. 206-208). Given the challenges of working with silicon, this was no easy task, and Teal knew it. While "most companies took the alloy route" (Goldstein, 1991) with respect to silicon due to the relative simplicity of the process, Teal chose to pursue grown junction silicon transistors. Ultimately, the alloy junction process was more a "matter of luck than anything else" (Goldstein, 1991, p. 27) as the silicon and junction materials expanded at different rates, causing substantial manufacturing challenges (Burgess, 2011b, p. 4). The grown junction process, on the other hand, was more "technique-oriented" and "controllable", where "time, temperature and pull rate pretty much determined the base width of the transistor" and ultimately the performance of the transistor (Wolff, 1985, p. 9). Teal's experience, knowledge and intuition about the advantages of the silicon grown junction method led TI to take a different path that, while challenging, allowed the firm to introduce the first commercial silicon junction transistor in 1954 ("Texas Instruments - 1954 first commercial silicon transistor," n.d.).

Additionally, Teal had become a bit of a celebrity within the industry and was able to recruit exceptionally talented scientists and engineers to join him at TI as he pursued the goal of silicon transistors. Mark Shepard noted that "we could never have attracted the stable of people that we did without him, or without somebody like him… and we got some really outstanding young scientists in those days." (Riordan & Hoddeson, 1997, p. 207) Teal used his contacts at universities to recruit some of the most promising engineers and scientists (Goldstein, 1991, p.

28). While many of TI's previous hires had been individuals with ties to Texas, such as Teal and Roger Webster (Goldstein, 1991, p. 4; Wolff, 1985, p. 4), Teal was able to recruit several individuals without previous ties to Texas who would prove instrumental in the production of the first silicon grown junction transistor, especially Willis Adcock from Standard Oil and Gas, Ross McDonald from Argonne National Laboratory, and Dr. Morton Jones from the California Institute of Technology (Goldstein, 1991, p. 31; Macdonald, 2009; "Willis Adcock - Wikipedia," 2011).

The superb talent and previous knowledge and experience with the grown junction process that TI possessed, however, did not result in the easy success some might have expected following Teal's move to TI. While he had successfully grown single crystal silicon with p-n junctions at Bell, he faced significant difficulties for more than a year at TI (Riordan & Hoddeson, 1997, p. 208). First, the lab had to be outfitted to perform the pulling operation. The equipment required for producing grown junction transistors had to be produced by TI, because there was no existing semiconductor production equipment industry at that time. This required substantial knowledge and skills that Teal was able to provide based on his Bell Labs experience (Wolff, 1985, p. 9). Following the outfitting of the lab, the group proceeded with the attempt to create transistors by pulling silicon and adding the necessary dopants. The pulling operation itself was very temperamental, and required a great deal of precision. However, the "careful selection" of temperature, rate, and amount of dopant was only possible following extensive experimentation to better understand the relationships between these aspects of the pulling process:

> "There were a great many crystals pulled in which they varied the time and rate of pull and the timing of the dropping of the doping pellets into the melt and that sort of thing in order to try and find a combination that would give a narrow base width and a fairly low base resistance... A lot of time and effort was spent in trying to develop a combination of pull rates, times and temperatures that could be used to get reproducible results." (Wolff, 1985, p. 11)

And even once the optimal values for temperature, pull rate, and dopant concentration were found, the process remained temperamental:

> "Crystal growing of semiconductor materials at that time was part science, part magic and a big dose of technique, with some prayers thrown in for added insurance. The crystal puller that we had at that time sported some dents in the quarter inch aluminum frame, with additional shoe and boot marks imprinted on it, evidence of earlier frustrating 'caresses' when it (we) didn't achieve the results we were hoping for." (Ward, 2001, p. 2)

These difficulties seem to have been driven by both a lack of knowledge among the staff about the pulling process as well as the purity of the silicon that TI was purchasing, both of which would eventually be overcome. After Teal used high purity silicon that he was able to source from DuPont for $500 per pound, he was able to successfully grow an NPN structure within silicon that was able to function successfully as a transistor. Within one month of the first successful growth of a silicon-based NPN structure, TI had established a production line and Teal made the announcement of the first silicon transistor at the 1954 IRE (Institute of Radio Engineers) conference in Dayton, Ohio (Riordan & Hoddeson, 1997, pp. 208-209).

Due to the higher melting point of silicon with respect to germanium, silicon transistors could operate reliably at temperatures and conditions that were not possible with germanium transistors, an advantage that was particularly valuable to defense firms (Riordan & Hoddeson, 1997, pp.207-209) who were willing to pay a premium for such devices. Following the introduction of the first commercial silicon transistor, TI enjoyed a monopoly in this market until 1958, with TI sales increasing "almost vertically" (Goldstein, 1991, p. 39). Ross Macdonald, former head of the Central Research Labs at Texas Instruments, indicated that "for quite a while nobody else could make them work well" (Kowalski, 2011a). It appears that there were three reasons for TI's monopoly. First, production of silicon grown junction transistors required a significant amount of knowledge that could only by gained through experience. Additionally, as explained below, although Bell Labs beat TI to producing the first silicon transistors, it ended up abandoning work on silicon grown junction transistors shortly thereafter. Finally, TI was very secretive about the information that it possessed, leaving other firms with no source for information on the silicon grown junction process after Bell Labs stopped pursuing this type of transistor. Only after the invention of the mesa transistor in 1958 did TI's monopoly end.

While one of the advantages of the silicon grown junction process as compared to the alloy junction process is that it is more controllable and "technique-oriented", this control and technique was something that had to be developed and learned. The overarching science describing the process was incomplete at the time, and thus the knowledge required had to be acquired through experience. As the previous quote referring to the indentations on the puller machine demonstrated, learning about the process involved trial and error, and even when the process had evolved to a state where transistors were produced, there was still significant art

remaining in the process (Ward, 2001, p. 2). In addition to the knowledge needed to perform the process, it should also be noted that a great deal of knowledge was required within the organization to build the required machinery for not just the pulling process, but also prerequisite processes as far back as production of pure silicon. All of these skills had been internalized by TI, to the point that eventually TI put DuPont, originally the only external supplier that many semiconductor firms used to source high purity silicon, out of the business of producing silicon (Burgess, 2011b, p. 27). Firms without knowledge and experience related to silicon production as well as the production of manufacturing equipment related to the various steps required to produce silicon grown junction transistors were at a serious disadvantage compared to TI.

Prior to TI's announcement of the silicon transistor at the Dayton IRE meeting, Bell had previously fabricated the first silicon transistor in January 1954. Morris Tanenbaum had joined Bell in 1952 and continued Teal's work with Buehler, Teal's former lab technician (Riordan, 2004, p. 2). Using the rate growing process, a variant of the grown junction process where both impurities are added to the molten bath at the same time, the Bell team was able to produce a silicon transistor that exhibited amplification. However, the team at Bell decided against patenting the process because of its similarity to previously used processes as well as a perceived lack of controllability of the process (Riordan, 2004, p. 3). As Tanenbaum noted later, "from a manufacturing point of view, it just didn't look attractive" (Tanenbaum, 2008, p. 3). Given the excitement around Bell regarding the new, promising technique to fabricate high frequency silicon transistors called diffusion, silicon grown junction transistors were abandoned, leaving TI as the only firm with the capabilities and knowledge to produce them (Lécuyer, 2006, p. 153).

While Bell had regularly shared "everything it knew" with other firms in the industry, the atmosphere at TI was very different. Rather than sharing and widely licensing, TI was very secretive about its work. One example of this secrecy was that while Teal wrote scientific papers to inform the community about basic discoveries at TI, some within the industry complained that Teal never "disclosed a lot of the details of the process to get the crystals to grow", leaving much of the needed knowledge to produce the crystals within TI (Burgess, 2011b, p. 5).

Additionally, while Teal became very active in the Dallas section of the IRE, he noted that TI was "only too glad to have people know what some of our accomplishments were", but the dissemination of information was selective, as "you wouldn't necessarily want to spread the news too fast" (Goldstein, 1991, pp. 51-52). Finally, another example of TI's secrecy comes from a gathering for industry participants at the IRE Solid-State Devices Research Conference in June 1954. Tanenbaum had presented his work at Bell on the silicon grown junction transistor, at which point "Teal mentioned similar work that had been done at TI – but was cagey about specifics" (Riordan, 2004, p. 3). Given that all three instances noted above involve Gordon Teal, it is unclear whether the secrecy practiced by Teal was indicative of company policy or simply Teal's persona. However, given the role that Teal played within TI, it seems that one would most likely have to go through him to gain access to the knowledge necessary to pursue silicon grown junction transistors. TI's secrecy, combined with Bell's abandonment of work on the silicon grown junction, resulted in a void with regards to a source where other firms could learn about the silicon grown junction process.

In summary, the silicon grown junction production method required substantial knowledge that was gained only through experience. While both Bell and TI had developed this knowledge, Bell gave up on the production technique shortly after it produced its first silicon transistor using the method. This abandonment by Bell cut off other semiconductor firms from the valuable information needed to produce using this method, which many of these firms would have traditionally accessed through their Bell Labs transistor license. Coupled with the secrecy of TI regarding this process and others utilized by the firm, other firms were unable to access the requisite knowledge and funding needed to pursue a program in silicon grown junction transistors.

While the explanation for TI's lengthy monopoly may revolve around the reasons listed above, the time at which the diffusion process became popular within the industry may have also played a role in ensuring TI's monopoly. Specifically, it appears that by 1955 the industry was endeared to diffusion, a new production process developed by Bell Labs, and in 1956 Bell was telling everything it knew to interested parties willing to pay for a license (Burgess, 2011b, p. 27). Given that it took TI, a firm with both previous crystal pulling experience and Gordon Teal, a co-inventor of the grown junction process, almost one and a half years to begin producing silicon transistors with the grown junction method, it seems unlikely that other firms without access to such knowledge and experience would be able to accomplish such a goal in the same time frame. One and a half years following TI's silicon transistor disclosure would be right in the middle of the industry's excitement over diffusion. At that point, perhaps firms took the same attitude as Bell that diffusion was the most promising production method of the future, and abandoned their pursuit of the silicon grown junction transistor.

Much like the germanium grown junction transistor, the silicon version exhibited less noise than the point junction transistor. But again due to the nature of the grown junction process, its frequency response was inferior to the point contact transistor due to the width of the base, and the device was prone to failures because its junctions were exposed. The biggest advantage of the silicon grown junction transistor, as mentioned before, is that it had consistent operation at much higher temperatures, which was important for military and space applications (Morris, 1990, p. 35). However, silicon grown junction transistors were simply "too expensive and difficult to make to be long term survivors" (Burgess, 2011b, p. 23) and were eventually replaced by more economical designs, including mesa and planar transistors.

Mesa Transistor

While the development of the silicon grown junction transistor created a device that was usable at higher temperatures, the lack of transistor devices that could operate at higher frequencies continued to frustrate the industry. High frequency response required a very thin base layer – something that had consistently eluded the industry (Moore, 1998, p. 36). However, the time had come to address this issue by combining two innovations from Bell Labs – double diffusion and oxide masking. While Bell Labs had disclosed the idea of a new type of transistor combining these two ideas, a resourceful team at a recently formed firm, Fairchild Semiconductor, would take the lead in developing this concept into a producible product. This new device, the mesa transistor, finally increased the operating frequency range of transistors beyond 50MHz, (Moore, 1998, p. 36) and was capable of operating at frequencies greater than 100MHz.

Working toward the goal of producing transistors with thinner base widths, the production method of diffusion seemed like a natural choice to develop new devices. Diffusion involves placing a semiconductor substrate inside of a furnace and exposing the substrate to a dopant, either p or n type, in order to introduce p or n regions within the semiconductor substrate. Varying temperature and time of exposure, the depth of each diffused region can be controlled precisely (Reid, 2001, pp. 88-89).[15]

Once Bell decided that it would replace all of its switching equipment in its central office with transistors in 1954, it became clear that transistors with accurate switching behavior were the top priority for Bell Labs (Riordan & Hoddeson, 1997, p. 221). This priority brought to the forefront one of the largest drawbacks of using germanium – it produces "leaky" switches like "a maddening faucet that you can never quite shut off completely" as the devices continue to "drip electrons" (Riordan & Hoddeson, 1997, p. 221). Given the importance of switches that turned completely off within the Bell system, the organization needed to pursue other semiconductor materials, namely silicon, which does not exhibit this leaky behavior (Riordan & Hoddeson, 1997, p. 221).

However, experimentation with silicon uncovered a weakness with that material as well. During the diffusion process, the surface of the silicon substrate was routinely damaged, resulting in a crystal that had been "eroded and pitted, or even totally destroyed" (Riordan & Hoddeson, 1997, p. 221). The Bell team struggled with this issue for months, when one day the individual

[15] In the early days of the industry, the diffusion process was much more an art than science, but subsequent experiments resulted in precise models of diffusion with respect to temperature and time.

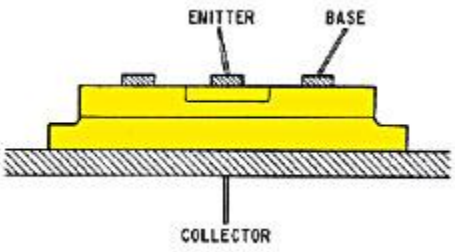
operating the diffusion chamber accidentally introduced water vapor into the chamber. This mistake caused silicon oxide to form on the surface of the substrate, which, instead of ruining the substrate as was commonly thought at the time, protected the surface from pitting during the diffusion process. The result was a perfectly useable diffused silicon wafer, and silicon's weakness had been addressed, enabling the researchers at Bell to move forward in developing diffused transistors with the material (Riordan & Hoddeson, 1997, p. 222). In fact, the development of silicon oxide as a masking agent would serve to fortify silicon's place as the preferred material for semiconductor development. The reason is that germanium oxide was unable to be used in the same way as silicon oxide because it is water soluble, causing it to be washed away during production processing (Riordan & Hoddeson, 1997, p. 222).

Having solved the silicon pitting problem, Bell researchers moved on to demonstrate double diffusion transistors. While diffusion had been used previously to construct alloy junction transistors (the substrate material was diffused to create a p region to be used as the base, and then separate regions were constructed by alloying the indium pellets), double diffusion uses several rounds of diffusion to create all three regions within the substrate (Riordan & Hoddeson, 1997, p. 223).

Combining the concepts of double diffusion and oxide masking by Bell created a new type of transistor – the mesa transistor. To produce these transistors, double diffusion was first used to create a PNP sandwich within the silicon substrate. In order to take advantage of the narrow base width, a portion of the p-n junction had to be removed. Otherwise, the base capacitance would be too large and the frequency range of the device would be diminished (Moore, 1998, p.

36). To accomplish this removal, the surface of the transistor was masked with an oxide, the windows within the oxide were then created, and the transistor was then etched to remove the necessary areas (Riordan & Hoddeson, 1997, p. 262). An illustration of the resulting mesa transistor, named for its shape, is shown below in Figure 7.

Figure 7: Mesa Transistor.



Source: http://spectrum.ieee.org/image/47335

While the mesa transistor was invented and developed at Bell Labs, Bell had yet to take the technology from lab to production. Some substantial issues in mass producing mesa transistors remained, but a new firm located in California – Fairchild – was determined to resolve these issues to create its first product.

Following the invention of the transistor, Shockley had grown less content with the situation at Bell. Shockley's ego, competitiveness and approach to managing people had alienated many that he had worked with, most particularly Bardeen, one of the co-inventors of the transistor (Riordan & Hoddeson, 1997, pp. 225-226). His professional advancement had been rather slow at Bell, and he felt that he was in effect "stuck" in middle management, which was not in line with his scientific and professional aspirations (Riordan, 2007, p. 36). After exploring several

career options, he decided to form his own firm, Shockley Semiconductor Labs, with the backing of his friend, industrialist Arnold Beckman (Riordan & Hoddeson, 1997, pp. 233-234). The firm was formed in Shockley's hometown, Palo Alto, CA, in 1955, and Shockley was able to recruit extremely talented and able employees from the East Coast[16] with the promise that the firm would be pursuing cutting-edge work in commercializing double-diffused transistors (Moore, 1998, p. 53). While early work at Shockley did involve silicon-based double diffused transistors, Shockley's focus turned to a four-layer diode, a device that Shockley had been enamored with since his days at Bell. Ideally, the four-layer diode could provide a much more advanced and reliable switch for networks such as Bell's phone network (Riordan & Hoddeson, 1997, p. 267). Shockley felt that if he could produce a product that would be purchased by Bell the success of his company would be guaranteed. This change in direction, as well as concerns with Shockley's management style, caused the "traitorous eight" to decide to start their own firm, Fairchild Semiconductor, with the financial backing of Fairchild Camera & Instrument (Bassett, 2002, p. 45). The arrangement was mutually beneficial: the founders got to "be their own boss" (Lécuyer, 2000, p. 164), and Fairchild Camera & Instrument was able to gain relatively quick entry into silicon semiconductors in exchange for initial funding (Lécuyer, 2000, p. 165).

While a quick read of history might lead one to think that Shockley's main contribution to Fairchild was in bringing the eight founders together, there appears to be much more to the story. In an article written for IEEE, Gordon Moore details the formative years of Fairchild and

[16] "The founding group of eight consisted of a metallurgist, S. Roberts; three physicists, J. Hoerni, J. Last, and R. Noyce; an electrical engineer, V. Grinich; an industrial engineer, E. Kleiner; a mechanical engineer, J. Blank;" and Gordon Moore, a physical chemist. (Moore, 1998, p. 53)

indicates that the group learned a great deal while at Shockley, much of which they would use when they chose to go out on their own (Moore, 1998, p. 54). Three examples seem to be particularly telling. First, the building that Shockley had selected for his firm had to be outfitted as a workplace for semiconductor research, design and manufacturing. Once this had been accomplished, the group realized that very few production machines (furnaces, etc.) were commercially available and thus had to be constructed in-house. Finally, the group was engaged in development and design work that taught them a great deal about silicon and semiconductor design. Moore notes that the employees at Shockley were performing research and development on the basic technology involved in double diffused transistors, and that the group worked to reproduce the diffusion and mesa transistor results that had been produced at Bell Labs (Lécuyer, 2000, p. 162). The work aimed at developing an understanding of the problems encountered through this replication process, because "neither the processing nor the physics of [silicon] was well understood" at the time (Lécuyer, 2000, p. 162).

The importance of training under an industry leader like Shockley cannot be understated as only one of the eight that founded Fairchild had experience with semiconductors prior to joining Shockley (Moore, 1998, p. 53).[17] The Fairchild founders appear to have recognized the significant knowledge and experience that they attained at Shockley. In a solicitation letter sent to potential investors, the group indicated that they had already "mastered the complicated techniques needed to produce semiconductors at their previous employment" (Lécuyer, 2000, p.

[17] The one founder with experience was Bob Noyce, whose experience was with germanium semiconductors (Moore, 1998, p. 53).

163). Taking the knowledge learned through these experiences with them, the group started Fairchild Semiconductor in 1957, just a few miles up the road from Shockley. The ultimate goal was to take the idea of the mesa transistor "and turn it into a product that could be manufactured reproducibly" (Moore, 1998, p. 53).

Bell Labs had described the design and technology and had produced limited runs of mesa transistors, and initial development and design work had been performed at Shockley, but the device was still far from being able to be mass produced (Lécuyer, 2000, p. 162). Before beginning serious development work, the Fairchild founders first had to create a work environment conducive to semiconductor research. This involved creating (to as much an extent as they could) a clean room environment and building both silicon growing and diffusion furnaces (which were not available for purchase from equipment manufacturers), both tasks that they had helped perform at Shockley (Moore, 1998, p. 54). Once the office and proper resources were in place, the group split up the tasks that required attention prior to the production of the device. As Moore recounts:

> "We divided the work to fit the backgrounds of the group. Roberts took responsibility for growing and slicing silicon crystals and for setting up a metallurgical analysis laboratory. Noyce and Last took on the [photo]lithography technology development, including mask making, wafer coating, exposure, development, and etching. Grinich set up electrical test equipment, consulted with the rest of the group on our electronic questions, and taught us how to measure various transistor parameters. Kleiner and Blank took charge of the facilities and set up a machine shop to make the equipment and fixtures we could not

> purchase. I took on the diffusion, metallization, and assembly technology development. Hoerni, our theoretician, sat at his desk and thought." (Moore, 1998, pp. 55-56)

The hurdles for the team were both numerous and novel. However, the team was able to overcome the major obstacles hindering the production of mesa transistors within five months of its first order from IBM through an intensive period of trial and error experimentation and innovation. These efforts would range from the growing of silicon wafers to the packaging of the transistors, with particularly notable efforts occurring in the areas of photolithography equipment, photoresists, diffusion, packaging, and defect testing.

Noyce and Last had tackled the development of photolithography, a process that produces templates for the areas of the semiconductor substrate that should be doped using diffusion and transfers those templates onto the substrate for further processing. While the process had been demonstrated by Bell Labs and the US Army's Diamond Ordnance Fuse Laboratories (Museum, n.d.), there was no commercially available equipment to perform photolithography on semiconductors. Therefore, Noyce and Last had to be resourceful in developing the process and equipment to be used at Fairchild. The two scoured San Francisco camera stores to gather three off-the-shelf lenses that had the most similar focal lengths in order to construct one of the first step-and-repeat cameras used to create the masks needed for photolithography (Moore, 1998, p. 56). Alignment of lenses in the step-and-repeat camera had to be exact in order to produce masks precise enough for manufacturing (Moore, 1998, p. 56). Noyce and Last developed an innovative array mounting method that maintained alignment through the mask creation process,

as well as a method to mount the masks using three reference points during diffusion to ensure proper alignment (Moore, 1998, p. 56).

Once the masks had been created and mounted, photoresist material was then deposited onto the semiconductor substrate. The photoresist, which had been deposited onto areas of the silicon surface that were to be protected from diffusion, would then be exposed so that it would harden and be able to protect the underlying substrate from dopants. However, existing photoresists would not work with silicon, as they would not adhere to the silicon oxide used in the manufacturing process. Additionally, existing photoresists introduced impurities into the silicon that led to major yield problems (Lécuyer & Brock, 2010, p. 20). Working with Eastman Kodak, which had originally developed photoresists used in commercial photography, Fairchild changed the makeup of the photoresists so that they would adhere properly to silicon oxide while not introducing impurities into the silicon (Lécuyer, 2000, p.169; Lécuyer & Brock, 2010, p. 20). While adherence issues with photoresists on semiconductors persist to the present day (Brock, 2006, p. 89), the advances made by Fairchild in cooperation with Eastman Kodak were substantial and allowed the firm to move forward in the production of silicon-based mesa transistors.

Problems in the diffusion process were also a significant hurdle that had to be overcome in the production process. Moore, having agreed to work on diffusion, noticed that the junctions he was creating were “soft”, meaning that they had poor electrical characteristics. Fairchild employees had seen this problem at Shockley and had some ideas as to how to resolve the issue (Moore, 1998, p. 56). After trying several unsuccessful approaches, Moore solved the diffusion

problems when he plated the back of the wafer in a process known as “gettering”. This process later became an industry standard in increasing the performance of junctions (Moore, 1998, p. 56). In addition to working on the specific problem of soft junctions, Fairchild also pursued research and development work to better understand the diffusion of dopants in order to engineer more controllable and economic diffusion techniques (Lécuyer, 2000, p. 169), which would result in products with a higher yield.

Not all of the innovations created at Fairchild were necessary to successfully produce the mesa transistors. Rather, some came through cost-cutting efforts or in a reaction to customer needs, and packaging innovations at Fairchild also followed this pattern. Initially, Fairchild used the same types of hermetically sealed metal cans that had been used for vacuum tubes on its transistor products (Lécuyer, 2000, pp. 169-170). Fairchild had implemented a key innovation with respect to packaging by mounting semiconductor chips directly to packaging containers, which resulted in greater robustness with respect to shock and vibration compared to the industry standard of connection through wires (Lécuyer, 2000, pp. 169-170). This innovation could be implemented, to an extent, regardless of what type of packaging was used, and the metal can used for packaging contributed a significant portion of the cost of the transistor. As a result, alternatives were sought within the organization to decrease both the cost of the component and the amount of labor required to use the transistor within other systems.

To answer this challenge, Fairchild developed plastic encapsulation, which replaced the costly metal can with a ceramic bead that the transistor was placed upon and then covered by epoxy. The materials required for the process were cheap, and the process was simple, allowing it to be

performed by low skilled (and low wage) workers (Lécuyer, 1999, p. 203). This innovation, along with other cost-cutting measures, decreased the cost of Fairchild's silicon transistors to a level that was competitive with germanium transistors and vacuum tubes (Lécuyer, 1999, p. 203).

Finally, the exposed nature of the junctions caused electric fields to form during operation, which attracted any impurities that were still located within the cans that transistors were packaged inside, and caused the transistors to malfunction. Fairchild employees first had to identify what was causing the malfunctions, and then devise a quality assurance process where the cans were tapped with a pencil in order to loosen all remaining impurities. Once this process had been performed, the transistors were powered up to determine whether they were of high enough quality to sell to customers (Moore, 1998, p. 57). However, this was not a long-term solution. Rather, the design of the transistor had to be examined and improved upon in order to create a more robust component. This would be accomplished with the advent of the planar transistor at Fairchild.

Fairchild employees were definitely able to use the knowledge and skills they had gained at Shockley to help them commercialize silicon mesa transistors. However, the innovations detailed above indicate that significant work was performed following the founding of Fairchild, leading to the first commercially produced mesa transistor. Still, this transistor had substantial yield and reliability issues that had to be improved in order to gain commercial acceptance (Lécuyer, 2000, p. 174). In fact, the push for increased reliability would mostly come from Fairchild's first major customers, namely IBM and Autonetics, while the push for increased yield

was internal and related to the unit costs. This customer desire, combined with significant know-how and resources that Autonetics would provide to Fairchild and the acquisition of talent from other skilled manufacturing firms would transform Fairchild from a startup semiconductor producer to a firm able to inexpensively (due to increased yields) produce reliable transistors.

Beginning with the first contract that Fairchild won, its customers were concerned with the reliability of its product. Sherman Fairchild, the founder of Fairchild Camera & Instrument, was also the son of the co-founder of IBM. While Fairchild Semiconductor had pitched its products to IBM separately, Sherman Fairchild played a role in convincing Thomas Watson Jr., IBM's CEO at the time, to purchase transistors from Fairchild Semiconductor. He told Watson that buying transistors from the new firm was "the safe thing to do" (Lécuyer, 2000, p. 168). As a part of IBM's agreement to purchase transistors from Fairchild, IBM engineers specified the desired electrical parameters and required Fairchild to use procedures specified by IBM in order to test the components (Lécuyer, 2000, p. 168). Fairchild's next major customer was Autonetics, the avionics division of North American Aviation and a producer of electronics for aviation systems. Autonetics had been awarded a government contract to manufacture the navigation and control systems for the Minuteman inter-continental ballistic missile program (Lécuyer & Brock, 2010, p. 23). Given the demands on the program for high-reliability, Autonetics agreed to purchase transistors from the firm "with the understanding that the company would improve the reliability of its devices by several orders of magnitude" (Lécuyer, 2000, p. 173). Luckily for Fairchild, Autonetics was willing to invest substantial resources in order to see the desired increase in device reliability, and the firm would play a major role in re-shaping Fairchild's manufacturing capabilities.

Following Autonetics' purchase agreement, Fairchild began implementing a "reliability improvement plan" to reinforce the company's manufacturing discipline, tighten process controls, and augment existing testing procedures (Lécuyer, 2000, p. 177). Fairchild was not the only Autonetics subcontractor targeted with such a program, but rather this quality improvement program was implemented at all Autonetics subcontractors (Lécuyer, 2000, p. 177). The plan required that Fairchild carefully document its manufacturing processes, create "high-reliability" production lines with dust-free environments (Lécuyer, 2000, p. 181), perform life testing on components, and implement processes to place serial numbers on each device to trace quality issues throughout the production system (Lécuyer, 2000, p. 178), among other requirements. Documentation was a key aspect of the improvement plan, and Fairchild developed a detailed production manual in cooperation with Autonetics. These manuals were hundreds of pages long and described the manufacturing process in excruciating detail. Workers were supervised to make sure that they were following protocol, and operators who were not were dismissed (Lécuyer, 2000, p. 170). Additionally, Autonetics was able to monitor the production facility at any time and step in to change processes as needed – a significant amount of control for a customer (Lécuyer, 2000, pp. 178-179). Describing the Autonetics program, Lécuyer notes:

> "Under the guidance and scrutiny of Autonetics, Fairchild's manufacturing engineers also perfected their production systems and tightened their control of the manufacturing process in order to produce highly reliable diodes and transistors" (Lécuyer, 2000, p. 181).

Ultimately, the reliability of Fairchild's products was substantially increased, and the firm was able to obtain higher yields and lower manufacturing costs. This system of manufacturing controls and improvements was later applied to planar products in a desire to achieve similar results (Lécuyer, 2000, p. 179).

While Autonetics applied direct influence on Fairchild's manufacturing systems and processes, industry practices also affected the Fairchild organization, as Fairchild both hired individuals from other firms with manufacturing experience and reacted to practices within the electronics industry. Early on, when the Fairchild founders hired Ewart Baldwin as the general manager, he brought a large group of manufacturing and instrumentation engineers with him from Hughes (Lécuyer, 2000, p. 167). At the time, Hughes was one of the largest producers of semiconductor diodes. The Hughes engineers were augmented by engineers from GE and Ford who were hired to improve Fairchild manufacturing processes (Lécuyer, 1999, p. 194). Charles Sporck, hired from GE, noted that there was "enormous pressure to cut costs" in manufacturing, and Sporck was able to apply aspects of the GE production system to create new organizations that "addressed problems as teams instead of adversaries" (Lécuyer, 1999, p. 194). In addition to engineering practices from Hughes, GE and Ford, Fairchild employed industry practices such as using skilled female workers and the creation of a pre-production engineering group to scale up manufacturing processes (Lécuyer, 2000, p. 170). Fairchild's commercialization of the mesa transistor in 1958 created a device that performed at significantly higher frequencies than other previous transistors (Moore, 1998, p. 57; Morris, 1990, p. 36). More important, the design made it easily mass-produced unlike previous transistors.

The mesa transistor manufacturing process made the batch production of transistors possible, as multiple transistor masks were able to be applied to a single wafer. Compared to alloy or grown junction production processes that only allowed for one transistor to be produced at a time, mesa transistors could be produced more efficiently, and thus at a lower piece cost (Braun & Macdonald, 1982, p. 74). However, the exposed junctions of mesa transistors caused significant reliability issues. Ultimately the long term fix for the reliability issue with the mesa transistor was the invention of the planar transistor, but prior to this advancement Fairchild implemented many practices to increase the reliability and yield of mesa transistors. These practices were motivated by customer desires and internal goals for costs, but they were developed through an Autonetics-mandated (and Autonetics-funded) program as well as by hiring individuals from other firms with mass-production manufacturing experience.

Planar Transistor

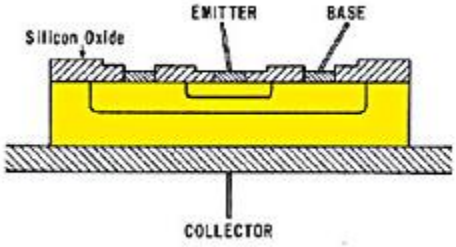
As the other founders at Fairchild were grappling with the production of Mesa transistors, Jean Hoerni was already thinking of a different type of transistor (Moore, 1998, p. 58). No more than two months after the founding of Fairchild, Hoerni had synthesized a design for a new type of transistor (Riordan, 2007, p. 39) that would revolutionize transistor production and ultimately lead to the creation of a new type of product all together – the integrated circuit. This new type of transistor was called the planar transistor, in reference to its flat surface. Additionally, the planar transistor would have characteristics unlike any transistor seen before, and its production process would build upon the knowledge and skills that were developed at Fairchild as the mesa transistor was commercialized.

Like many other types of transistor designs, the planar transistor emerged because of deficiencies its predecessors. Fairchild, having signed an agreement with Autonetics to provide transistors to them for the Minuteman program, found itself unable to produce transistors that were reliable enough for the application (Lécuyer, 2006, 149). After employees determined that loose particles inside of the transistor packaging were to blame for failures, significant effort was directed at addressing the problem by cleaning up the manufacturing process and designing new test procedures (Lécuyer, 2006, 149). However, these efforts were unable to resolve the reliability problems with the mesa transistors being produced by Fairchild. Marketing was placing significant pressure on engineering to "solve this damn tap test problem" (Lécuyer & Brock, 2010, p. 153) and Jean Hoerni would eventually succeed in solving the problem by building on Fairchild's previous experience with silicon oxide in the development of the mesa transistor and introducing what was then known as a "dirty" material into the mix.

In spite of the full implications of the Frosh and Derick paper, most in the industry limited the use of the oxide to a masking agent during fabrication, and it had become standard industry practice to remove the "dirty" oxide that was created during fabrication. Contrary to industry belief, Jean Hoerni of Fairchild believed that the oxide might be retained following fabrication to protect the silicon surface (Moore, 1998, p. 58). The key to the reliability problems that were being experienced with mesa transistors was having an unprotected silicon surface, and an oxide layer was one possible way to protect the surface. With this in mind, Hoerni began to test his hypothesis that oxide protection would lead to a more stable, reliable transistor.

Hoerni's innovative insight with what became known as the planar process was in a large part related to the advances of photolithography that Fairchild had made during its pursuit of the mesa transistor. While planar fabrication used many of the same skills and methods (photolithography, diffusion, etc.), there remained significant differences in the production of the two types of devices. There were three main differences between planar transistor production and mesa transistor production. First, there was no etching of exterior areas of the transistor and once diffusion operations were completed a final layer of oxide was added to the transistor. Second, the location of metal leads to the regions was removed in order to allow for successful binding of the leads to the transistor (Morris, 1990, p. 38). And finally, one extra mask had to be produced to support planar fabrication as compared to mesa fabrication (Moore, 1998, p. 58). Initially, Hoerni fabricated a partial planar transistor, only protecting the emitter-base junction. The transistors produced in this fashion had "higher gain… and were more electrically stable than conventional mesa transistors" (Lécuyer & Brock, 2010, p. 141), providing support to the idea that the planar process would produce more reliable devices. This evidence encouraged Hoerni to pursue full planar transistors that had oxide protection over both the emitter-base and base-collector junctions. An illustration of a planar transistor is shown below in Figure 8.

Figure 8: Planar Transistor.



Source: http://spectrum.ieee.org/image/47335

The resulting transistor was unlike any other transistor previously produced. Not only was the gain (amplification) much better than mesa transistors, but the leakage current was significantly less than previous transistors. Hoerni noted, "I remember I started to change the units I was using to measure leakage counts because they were a thousand times less" (Lécuyer, 2006, p. 152). Most notable, in terms of improvements, was the reliability of planar transistors. To demonstrate the increased reliability of planar transistors, Hoerni performed a dramatic demonstration when first sharing his results with the group at Fairchild. After banging on the transistors with a hammer, he ran them through the tapping test machine that had been used for mesa transistors and proceeded to demonstrate that the devices worked flawlessly (Lécuyer & Brock, 2010, p. 31). As Hoerni noted, "the most interesting thing [with these devices] was once they were sealed then you could tap forever, [and] nothing would happen" (Lécuyer, 2006, p. 152). This type of reliability had never been seen before in transistors and would be vital in increasing the usage of semiconductors in the future. Additionally, the planar process provided improved batch processing opportunities; costs came down, and characteristics became much more reproducible, decreasing variability between transistors (Morris, 1990, p. 38). Finally,

since the main tool used in creating the transistor was the photolithographic mask, it was very easy to "retool" for different transistor variations (Morris, 1990, p. 38).

While the advantages of the planar transistor were impressive, there were substantial issues with regard to yield that caused difficulty in the commercialization of the planar transistor. Early planar transistor yield "did not exceed 5 percent", notably lower than the early yields of 25-30 percent for mesa transistors (Lécuyer, 2006, p. 152). Interestingly, Bell Labs had invented a process similar to Hoerni's, but "decided to forego its further development because of its seeming lack of manufacturability" (Lécuyer, 2006, p. 153). Fairchild, however, did not have the luxury of shrugging off the new type of transistor because it was facing increasing pressure to produce high-reliability transistors. The low yield rates of early planar transistors were caused primarily by poor oxide quality during fabrication. This deficiency allowed for dopants to diffuse into undesired locations of the wafer, causing significant performance issues (Lécuyer, 2006, p. 154). Over the year following the first disclosure of the planar process by Hoerni, Fairchild's research lab "devoted substantial efforts to understanding this complex process" (Lécuyer, 2006, p. 154) and fixed the problem by developing "oxidation techniques which enabled the growth of cleaner and more uniform oxide films on top of the silicon wafers" (Lécuyer, 2006, p. 154). Once the oxide layers became more consistent and dust and other particles were eliminated from various processes, yields similar to mesa transistors were achievable using the new planar process (Moore, 1998, pp. 58-59).

Technical challenges, however, were not the only type of challenges that Fairchild had to overcome in order to commercialize planar transistors. Both Raytheon and Hughes Aircraft

challenged Fairchild's planar process patent. Raytheon claimed that a previously issued Bell patent had significant similarities to Hoerni's method, while Hughes claimed that Hoerni's claims were anticipated by Hughes engineers (Lécuyer & Brock, 2010, p. 145). In what might be considered an effort to decrease the delays to licensing the planar process, both challenges were settled out of court, with Hughes receiving a portion of future royalties on the Hoerni planar process patent and Raytheon abandoning its case against Fairchild. It seems that the incentive to quickly settle the challenges was warranted, as Fairchild Semiconductor received more than $100 Million in royalties on both the planar patent and the integrated circuit in the following decade (Lécuyer & Brock, 2010, p. 145).

Once all technical and legal challenges were overcome, Fairchild was able to introduce the planar transistor to the market and begin mass production. Even at the time of introduction in 1960, the implications of the planar transistor were clear to many within the industry. One attendee at the conference where the planar transistor was introduced made these remarks regarding the presentation:

> "Fairchild… presented graphs… which were so much better than anyone had seen before that it was quite obvious that if they were genuine a real breakthrough had been achieved. After several hours' discussion with Grinich it became clear to me that the planar process was the process of the future." (Morris, 1990 p. 37)

And later, industry observers noted the broad impact of the planar transistor on the rest of the industry, by noting that "with the exception of certain applications, such as high voltage rectifiers and thyristors, the planar technique rendered all previous methods of device construction

obsolete." (Morris, 1990, p. 37) The planar transistor dominated most transistors for the vast majority of applications, and through its reliance on silicon oxide secured silicon's place as the preferred material for transistor manufacturing. Additionally, the advances in oxide masking and surface stabilization as well as metal-based interconnections originally introduced with the planar transistor would play important roles in the invention of the integrated circuit.

The planar transistor was clearly a large step forward in terms of device performance, reliability, and manufacturing capabilities. However, the question remains – why was Fairchild, arguably a small start-up firm, the firm to invent and commercialize this innovation when so many electronics producers had the chance to pursue this opportunity? In short, Fairchild had the motivation, skills, and flexibility.

In terms of motivation, Fairchild had motivation from its key customer, Autonetics, as well as competing firms, both existing producers and new startups. Fairchild had recently entered an agreement with Autonetics that required the firm to produce high-reliability transistors. While the firm believed that mesa transistors would meet the requirements for Autonetics, the latter shortly realized that this was not the case. Given the status of the Minuteman project as the largest defense program of the era ("Computer History Museum - The Silicon Engine | 1958 - Silicon Mesa Transistors Enter Commercial Production," n.d.), Fairchild was desperate to find a way to produce reliable transistors in order to salvage its production contract with Autonetics (Lécuyer, 2006, p. 153). Following Fairchild's disclosure of the planar process, Autonetics expressed significant interest in the new innovation, requested samples, and encouraged Fairchild to pursue planar transistors (Lécuyer & Brock, 2010, p. 144). Additionally, Fairchild

faced increasing competition, both from other existing firms, such as Motorola and TI, which had introduced mesa copies in late 1959 (Lécuyer, 2006, p. 153), as well as a new semiconductor operation named Rheem Semiconductor. Shortly before Hoerni disclosed the planar process to the management team at Fairchild, the firm's General Manager, Ewart Baldwin, decided to leave Fairchild and form Rheem Semiconductor, taking many key employees that he had brought with him from Hughes, as well as copies of Fairchild's process manual (Lécuyer & Brock, 2010, p. 162). Baldwin's objective was to produce cheaper versions of mesa transistors to compete with Fairchild. Given that Baldwin had all of the production information needed to reproduce Fairchild's product, the management team was very motivated to pursue Baldwin and Rheem through legal means[18]. Not knowing the outcome of the legal actions taken by Fairchild, the firm's management decided to make mesa transistors obsolete by fully developing planar transistors (Lécuyer & Brock, 2010, pp. 162-167).

The planar process was radical in terms of its approach to using silicon oxide, but the vast majority of processes were adopted from the mesa transistor production that had been occurring at Fairchild (Moore, 1998, p. 58). Arguably, with perhaps the exception of Rheem, there was no other firm that was as capable in terms of the needed skills and processes as Fairchild because of Fairchild's role in developing the mesa process. Additionally, Fairchild had maintained its flexibility regarding production. While other established firms, such as TI, Motorola, and Philco, were "far into major efforts in the 'automation' and 'mechanization' of transistor production" in order to lower manufacturing costs, Fairchild resisted any incentive to mechanize. By not

[18] The lawsuits against Baldwin and Rheem were settled out of court (Lécuyer & Brock, 2010, p. 162).

mechanizing, which would have locked the firm into a set production process, Fairchild was able to be more flexible than its competitors, which was "one of the factors which have contributed to Fairchild's prodigious success" (Lécuyer & Brock, 2010, p. 222).

Ultimately, Fairchild's motivation, skills and flexibility played a large role in Fairchild's leadership role in silicon transistors. This leadership role, as well as internal knowledge and skills, provided the firm with a competitive advantage as the complexity of production processes continued to increase with the invention of the integrated circuit.

Summary of Transistor Innovations

A table encapsulating the transistor families is shown below (Table 1). In addition to information about the firm responsible for originally producing the transistor family, the table summarizes advantages and disadvantages of each transistor family as well as the gains of each device.

Table 1: Comparison of Major Transistor Types, 1947-1962.

Transistor Type	Originating Organization	First Commercial Production	Advantages & Disadvantages	Gain
Point Contact	Bell Labs	1951	50MHz Frequency Response limit, Fragile	20-30
Germanium Grown Junction	Bell Labs & GE	1951	Decreased Noise, Limited Frequency Response (1-10MHz), Low Power Capability, Exposure of Junctions	30-50
Alloy Junction	GE	1952	Adaptable for Large Scale Production, Increased Frequency Response (5-10MHz), Exposure of Junctions, Low Power Output	30-80
Surface Barrier	Philco	1954	Increased Frequency Response (50MHz), Low Yields, Mechanically Fragile, Similar to Alloy Junction	30-80
Silicon Grown Junction	Texas Instruments	1954	Unpassivated surface, Limited Frequency Response, Higher Temperature Operation	9-40
Mesa	Fairchild	1958	Highly Accurate Control of Base Width, Increased Frequency Response (>100 MHz), Low Power Handling Capacity, Poor Switching Characteristics, High Breakdown Voltage	10-50
Planar	Fairchild	1960	Protected Junctions, Increased Reliability, Increased Batch Production Opportunities, Not Available for Germanium, Poor Power Capability,	100-800

Sources: Braun & Macdonald, 1982; Doremus, 1952; Morris, 1990; "Physical Fabrication of Transistors," n.d.; Riordan & Hoddeson, 1997; "Texas Instruments - Transistor History," n.d.; Tilton, 1971; "Vintage Semiconductors Ltd Transistors," n.d.

The development of transistor technology, while pioneered by Bell Labs, was pursued by multiple firms. With the exception of some of the early types of transistors, for which firms

"invented around" Bell's possible patent position, transistor developments can be tied directly to research and development work performed at Bell Labs, speaking volumes to the role played by the Federal government-mandated openness of the organization in disseminating its work on transistors. While the industry struggled for almost 10 years with transistor reliability problems tied to the unprotected surface of the device, this problem was overcome with the advent of the planar transistor by Fairchild in 1960. This design innovation yielded stable devices that could be used in high-reliability applications, and it would also provide inspiration for the invention of the integrated circuit. As discussed later, in the next era of the semiconductor industry, the openness that was practiced by Bell vis-à-vis transistors would not be found among any of the leading firms in integrated circuits. This shift in openness led to a very different industry dynamic with respect to knowledge flow and availability, resulting in a very different story of industry development.

Integrated Circuits: Motivation and Paths to Miniaturization

As semiconductor technology evolved throughout the 1950's, systems were becoming more and more complex due to the evolution of smaller components. This increasing system complexity led to a problem for the industry – the "tyranny of numbers". Put simply, electronic systems were the result of connections between each of the components of the system, and the increasing complexity of these systems increased the number of connections required in what typically was a circuit that was decreasing in size. Also, as each new component was added to a system, additional points with failure probabilities were introduced into the system. Because any new component added was discrete and had its own distinct failure probability, the increasing complexity of electronic systems resulted in systems that had decreasing reliability. One statistic

regarding electronics reliability was that, in the 1940's, half of the military's shipboard electronic equipment was down at any given time (Evans, 1998, p. 330). While the use of transistors in place of vacuum tubes increased the reliability of the system, there was a realization within the industry that reliability gains would plateau because of the nature of the interconnections between components. The increasingly complex nature of circuits required more and more connections between components, and each connection was a possible source of failure for the system (Reid, 2001, p. 16). The reliability problem, as well as a perceived limit to system complexity that could be achieved with traditional discrete components, drove the military and others within the semiconductor industry to pursue alternative methods for producing electronic circuits, often referred to as miniaturization, during the 1950's. This drive ultimately resulted in what is today called the integrated circuit.

One account of the invention of the integrated circuit at TI presents the motivation and thinking within the industry at the time:

> "What caused Jack Kilby to think along the lines that eventually resulted in a working integrated circuit? Kilby, like all engineers, was a problem-solver. During the 1950s, the electronics industry was grappling with a problem, aptly called "the tyranny of numbers" by engineers whose designs had been thwarted by its barriers. As electronics had become more sophisticated, engineers were able to design ever more complex electronic circuits and equipment containing hundreds or thousands of discrete components (such as transistors, diodes, rectifiers, and capacitors). These components had to be interconnected to form electronic circuits. Making these components and the connections

> in a cost-effective, reliable way presented great difficulties. Kilby [at Texas Instruments] and others were seeking a solution to this problem." (Merryman, 1988, p. 5)

Today, the integrated circuit story is often told beginning with the co-invention of the device by Fairchild and Texas Instruments (Reid, 2001, p. 115). However, the push for circuit miniaturization fueled broad technological efforts and developments for almost an entire decade before this occurrence.

Miniaturization Paths

The nature of the various miniaturization paths was in part based on the pre-existing skills and competencies of the firms that performed the development, as well as the needs of the parties who funded and supported this work. As a result, the miniaturization options varied significantly. Discussing the state of semiconductors in the early 1950's, Daniel Holbrook described it this way:

> "Semiconductor technology, in short, was in an uncertain state, and the path to greater certainty had many branches". (Holbrook, 1999, p. 150)

In addition to the needs of the various parties that were potential users of the miniaturized circuits, the uncertainty regarding the technology and industry also contributed to the number of miniaturization paths. The pioneering work in miniaturization was performed mostly by firms that had previous electronics experience. This may have conditioned the types of miniaturization efforts that were pursued by the industry. S.M. Stuhlbarg of electronic components manufacturer P.R. Malloy & Co. compiled an illustrative listing of firms that were pursuing circuit

miniaturization in 1961, which is shown below in Table 2. Firms with previous electronics experience are denoted with an asterisk, which account for the vast majority of firms involved in miniaturization efforts.

Table 2: Firms pursuing circuit miniaturization in 1961

Firm	Miniaturization Path(s) Pursued
*AMP Inc.	Hybrid
* Bendix Corp. Radio Division	Hybrid, Molecular
* Burroughs Corp.	Hybrid
Cleveland Metal Specialties Inc.	Modular
Corning	Film
Fairchild Semiconductor Corp	Monolithic
Francis Associates with Sippican Corp.	Hybrid
* General Electric	Modular
Halex	Film
* Hughes Aircraft	Modular
* Hi Q Div. Aerovox Corp	Modular/Film
* International Resistance Co.	Film
* IBM	Film, Hybrid
Litton Industries	Hybrid
Martin Co.	Molecular
* Motorola	Film, Molecular, Hybrid
* P.R. Mallory Co.	Modular
Radiation	Molecular
* Raytheon Manufacturing Co.	Hybrid
Republic Aircraft Corp.	Hybrid
* RCA	Modular, Monolithic
Servomechanisms	Film
Stupakoff	Film
* Sprague Electric Co.	Film
* Sylvania Products Corp.	Modular
* Texas Instruments	Monolithic
* Varo Mfg. Co.	Film, Molecular
* Westinghouse	Molecular

Sources: Holbrook, 1999; Reid, 2001; Stuhlbarg, 1961. Firm background information from author's compilation of firm histories and production records.

There were a large number of approaches to circuit miniaturization, from those closely related to existing technology, to entirely new paradigms with regard to electronics. By the early 1960's most approaches fell into five specific miniaturization paths that will be detailed in this section. It is important to note that any of these five miniaturization paths could be considered integrated

circuits in that the components for a circuit were all being integrated into a single device. Today, the term integrated circuit has a very specific connotation, namely that the device is composed on a single semiconductor substrate. However, this understanding of the term is conditioned on several decades of technological developments. For the purpose of this dissertation, the term integrated circuit should include any efforts that place multiple components (or functions) within a single device, which includes all of the five efforts described below.

Hybrid Circuits

The most conservative circuit miniaturization option from a technological perspective was called hybrid circuits. Generally, hybrid circuits combined discrete components[19] on a single circuit board, which typically had printed leads to connect the discrete components. While this approach may seem technologically basic, it is important to note that until this point the vast majority of circuits were designed by firms with in-house circuit designers, as opposed to being sourced from other firms (Kowalski, 2011b, p. 5; Lécuyer, 2006, p. 223). The use of hybrid circuits shifted that design task to hybrid circuit manufacturers, allowing systems firms to focus on how to integrate these circuits into the products they were manufacturing instead of designing and building the individual circuits from scratch.

Compared to other miniaturization options, the technical difficulty of producing hybrid circuits was relatively low. Rather than needing to invent various technologies in order to bring the components together, hybrid circuit producers were essentially assembling circuits – with

[19] Hybrid circuits typically combined both active (i.e. transistors) and passive (i.e. resistors) components.

increasingly automated processes – using standard off-the-shelf components. Additionally, the cost to produce these circuits was relatively similar to the cost associated with purchasing the individual components and assembling them into a circuit. While early in the industry's history this made hybrid circuits some of the most cost-competitive types of miniature circuits, cost decreases of other types of miniature circuits would cause hybrid circuits to fall out of favor for all but a few specific applications. These applications included linear circuits in the early era of integrated circuits, as well as components such as very large capacitors, wound components and oscillation crystals ("Hybrid integrated circuit - Wikipedia, the free encyclopedia," n.d.).

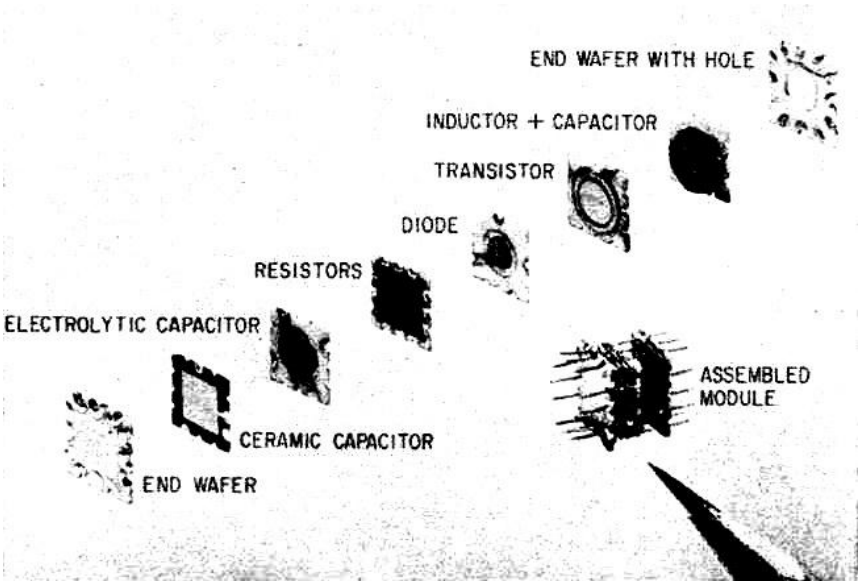
Modular Circuits

As previously mentioned, reliability had become a large concern with the increasing complexity of electronic systems. While the ultimate goal was to make electronic systems more reliable, another approach was to make the system much easier to service in order to replace failed components and thus increase overall up-time associated with a system. This was the concept behind modular circuits. Modularization was implemented in such a way that servicing circuits could be made easier for field applications by placing one or more components on a wafer that could be removed from the circuit easily. Wafers would be stacked on top of each other and then interconnected to form a circuit that would then be encapsulated for protection.

The most prominent modular circuit development program was Micro Module, a multi-million dollar program funded by the US Army Signal Corps. The desired outcome of the program was the creation of electronics hardware that could be easily interconnected in a way that would allow for smaller, more reliable, more serviceable hardware that could be used by the military.

By the early 1960's, the Micro Module program had succeeded in creating microelements that included resistors, capacitors, inductors, transistors, and diodes. An image of a typical Micro Module assembly is shown below in Figure 9. The program's prime contractor was RCA, which developed the overall framework for the technology and some of the modules. Yet, RCA also relied on some 60 sub-contractors that developed various component wafers, including Texas Instruments ("The Chip that Jack Built," n.d., "The Hybrid Microcircuit Micromodule & Solid Logic Technology," n.d.).

Figure 9: Sub-Assemblies of the RCA Micro Module.



Source: http://www.chipsetc.com/uploads/1/2/4/4/1244189/7461842.jpg?307

Although Micro Module was not the only modular circuit development program, it was the most widely known and involved the largest number of firms. Other firms active in modular circuit development programs included General Electric, Hughes, Aerovox, Mallory and Sylvania. Due to the nature of the modular pieces, most types of components could be implemented through

modular circuits. Additionally, the modules were easily serviceable, accomplishing the basic goal of the Micro Module program. However, the distributed nature of the component production, as well as the modular infrastructure required for each component, made modular circuitry more expensive than other approaches ("The Hybrid Microcircuit Micromodule and Solid Logic Technology," n.d.). By the mid 1960's, the Micro Module program had been abandoned as monolithic integrated circuits provided a more robust and economical approach to circuit miniaturization ("The Hybrid Microcircuit Micromodule and Solid Logic Technology," n.d.).

Film Circuits

Unlike modular circuits, film circuits used a single surface for the placement of circuit components. Passive components such as conductors, resistors and capacitors were formed by sputtering or evaporating conductive film onto a substrate, with active components such as transistors added later in their traditional form. The film was either deposited through sputtering, the practice for thin film circuits, or through screen printing the film onto the substrate, as was done for thick film circuits. Film circuits are considered by some to be a subset of hybrid circuits, with the key difference being that passive components are formed on top of a substrate as opposed to coming from off-the-shelf components. These circuits took advantage of a great deal of existing technology in a similar fashion to hybrid circuits. While firms producing film circuits had to invest in technologies and knowledge related to film sputtering and screen printing, these firms were also able to utilize many existing off-the-shelf components and processes.

Advantages of film circuits include the ability to include passive components with wider ranges (i.e., capacitance or resistance) and tolerances as well as better high frequency performance than other types of circuits ("Classification of Integrated Circuits," n.d.). This made film circuits ideal for linear or analog circuits, among other applications (Kabaservice, 1978, p. 63). However, circuit designs utilizing film circuits were typically more expensive and larger than those implemented with other miniaturization approaches (*Basic Electrical And Electronics Engineering (PTU, Jalandhar)*, 2006, p. 643). Therefore, hybrid circuits were preferred in many applications as compared to film circuits. Pioneering firms in film circuits included Varo, IBM, Motorola, and Servomechanisms.

Monolithic Circuits

Monolithic circuitry took a substantial step away from existing technology by creating all of the components of the circuit within a single semiconductor substrate (as opposed to placing components on the surface of the substrate). All components, passive and active, were formed by diffusing materials into the semiconductor substrate. Once the components were formed within the substrate, they would then be connected by overlaying interconnect material on top of the substrate. Firms that pursued monolithic circuits had to invest heavily in techniques to diffuse materials precisely into semiconductors and methods to protect the surface of the semiconductor from electrical shorts.

While monolithic circuits tended to be more expensive than equivalent hybrid or film circuits, monolithic circuits had the advantage of being able to incorporate more complex circuits since the interconnection process used for monolithic circuits was able to eliminate the tyranny of

numbers concern. Additionally, the integration of components within monolithic integrated circuits resulted in power requirements that were orders of magnitude lower than other types of circuits (hybrid, modular, etc.) ("Computer History Museum - The Silicon Engine | 1962 - Aerospace systems are first the applications for ICs in computers," n.d.). However, monolithic circuit production was a significant break from the norms associated with existing semiconductor component production, requiring significant investment and learning on the part of firms interested in pursuing the technology. Pioneering firms in monolithic circuits included Fairchild and Texas Instruments. These firms will both be detailed later in an effort to convey the extent to which firms interested in monolithic circuits in the early days of the technology had to invest in processes and technology to support production.

Molecular Circuits

Molecular circuitry was not based on existing electronic parts and methods. Rather, it focused on the molecular construction of crystals that would perform specific functions as opposed to the creation of circuits using traditional circuit elements (Kilby, 1976, p. 649). The approach promised that interconnects between components would not be needed, thus solving the "tyranny of numbers" problem because an entire semiconductor chip would represent the combination of many components. In theory, molecular circuits would be the ideal miniaturized circuit – one that would require no interconnections and thus one capable of performing the most complex operations with excellent reliability. One concern with this approach, however, is that conventional circuit design concepts could not be used with molecular circuits. Circuit designers had difficulty examining a circuit block that performed a given function as opposed to doing traditional circuit analysis by examining the individual components (Kowalski, 2011b, p. 11).

As a result, working with a molecular circuit without any real representation of traditional components proved to be quite a challenge for many industry players.

By the early 1960's, Westinghouse had developed 18 different types of functional blocks, but by 1962, had "altered its use of the term [molecular circuits] to mean monolithic integrated semiconductor circuitry" (Holbrook, 1995, p. 154). The molecular effort, having received extensive Air Force support since its inception, was abandoned in the early 1960's in favor of monolithic integrated circuits. Pioneering firms in molecular circuits included Westinghouse, Bell Labs, Bendix and Motorola.

Hypotheses: Integrated Circuits Entry and Performance

While there was significant interest in the development of integrated circuits to solve the reliability problems of the discrete component circuit, there was also little clarity as to what approach would best serve the industry. There was intense activity among firms pursuing the integrated circuit, and most firms were only active in one of the five miniaturization paths, leaving them at a potential disadvantage when one of the miniaturization paths would become most favorable for integrated circuit development. Additionally, the initial actors of the integrated circuit era were numerous, located within several key electronics regions. However, how does the actual development of the industry during the integrated circuit era compare with existing theories regarding firm entry and performance? What role did the existence of information within leading firms play in the development of the industry, and which theories best describe predictions that match reality? And how does the development of the industry in the early stages of the integrated circuit era compare with the early stages of the transistor era? This

section generates hypotheses using many of the general theories discussed in Chapter 2 in the context of what is known about the development of the integrated circuit.

When examining miniaturization paths, it is clear that, with the exception of molecular circuits, most miniaturization approaches used either conventional electronics or design procedures that were based on conventional semiconductor and electronic components. As a result, it seems reasonable to expect that firms previously producing semiconductor and other electronic components would have the skillset necessary to begin producing integrated circuits using any of these miniaturization paths. From Table 2, it is notable that the vast majority of firms that were active early in miniaturization efforts were incumbent electronics firms. This could be driven by two possible explanations: these firms felt compelled to develop integrated circuits because they were concerned that their products would be made obsolete by them, or these firms had the necessary capabilities to perform the initial research and development for integrated circuits.

The story regarding early entry into integrated circuits appears to be fairly similar to the scenario of entry into transistors. Like transistors in relation to tubes, integrated circuits had the potential to change the industry by moving the function of circuit design into circuit manufacturing firms from where it had previously resided, i.e., within the original equipment manufacturers. Up until this point, the original equipment manufacturers had been the customers of component manufacturers, selecting components to design into their own circuits. With the advent of the

integrated circuit, components[20] would be purchased in large part by integrated circuit manufacturers that would then design and manufacture circuits for original equipment manufacturers (Kowalski, 2011b). This created an obsolescence fear for incumbents that might have played a particularly large role in entry decisions.

Even if the concern regarding obsolescence did not induce firms to begin producing integrated circuits, it seems that related experience would have provided an advantage to component firms with regard to exploring opportunities associated with the production of integrated circuits, which were rapidly gaining importance. As noted previously, such related technical experience has been shown to positively affect a firm's propensity to enter other industries (Klepper & Simons, 2000; Lane, 1989; Scott-Morton, 1999). There were similarities between the production of electronic components, in particular semiconductor components such as planar transistors, and integrated circuits. These similarities would have been likely to provide an incentive for electronics firms to pursue integrated circuit production.

Regardless of which mechanism was at play, a higher propensity to enter integrated circuits among firms with related experience should be observed. Given detailed firm background information, the influence of previous experience can be tested with the following hypothesis:

[20] Discrete components were required to produce several types of integrated circuits, including hybrid and film integrated circuits. Clearly in the case of molecular and monolithic integrated circuits discrete components were not required, providing an even larger incentive for component manufacturers to begin producing integrated circuits in order to avoid complete obsolescence.

Hypothesis 7: Among existing firms at the time that integrated circuits were invented, firms with related electronics experience were more likely to begin producing integrated circuits than firms without electronics experience, all else held constant.

Which existing firms were most likely to begin producing integrated circuits has been explored. Yet, it is equally important to reflect on what specific firm traits (location, experience, etc.) might have allowed firms to perform particularly well following entry. Given the large number of diversifying electronic firms that were involved in pioneering integrated circuit development (see Table 2), many initial integrated circuit producers were located within the traditional electronics clusters previously identified: Boston, Los Angeles and New York. As detailed when discussing transistor producer performance, informal knowledge sharing, collaboration efforts and movements of workers between co-located firms are expected to have a positive effect on the performance of firms located within semiconductor clusters. Therefore, similar to what was hypothesized for transistors, one would expect that firms located within semiconductor clusters would perform better than similar firms located outside of these clusters.

Unlike the role played by Bell Labs in the transistor era, there was no unique source of cutting-edge technological information for all firms to gain knowledge from. Rather, the critical knowledge existed within a number of firms, as a large number of diversifying electronics firms located in the traditional electronics clusters were pursuing a variety of approaches towards integrated circuits. It seems likely that the dispersed nature of the knowledge would lend itself more easily to collaborations among firms and individuals that were co-located, or the movement of workers among these. If this was the case, co-location mechanisms might have played a key

role in the dissemination of information throughout the industry, especially in contrast to the emergence of transistors.

Silicon Valley was not previously a region of significance for transistor development or production. Yet, as noted in the previous section and further detailed in future sections, Fairchild Semiconductor emerged as a pioneering integrated circuit producer. Not only was Fairchild a pioneering firm, but it was also the source of a great deal of knowledge regarding the production of integrated circuits through extensive research and development efforts. This was especially relevant for monolithic integrated circuits, which would eventually become the dominant technology in the industry. Given that Fairchild was located in Silicon Valley, it seems reasonable to expect that firms co-located within Silicon Valley may be able to gain access to valuable integrated circuit knowledge through collaborations with Fairchild. Using performance data for integrated circuit producers, the presence (and effects of) regional agglomeration economies within semiconductor clusters during the integrated circuit era can be examined by testing the following hypothesis:

Hypothesis 8: Among integrated circuit producers, firms located in Boston, Los Angeles, New York and Silicon Valley are more likely to out-perform comparable firms located elsewhere, all else held constant.

When firms make decisions regarding the location of new operations, it seems reasonable to expect these firms to utilize their existing workforce and capital investments in such a way that new operations are likely to be located near existing operations. The firms know the region(s) around these operations well, reducing search costs involved with site location, hiring, sourcing of inputs, etc. However, this statement assumes that there are no advantages related to locating

operations in specific regions, such as decreased input costs, labor pooling, and access to specific knowledge existing within the region. If these advantages exist within a specific region(s) due to agglomeration and can compensate for potentially higher costs related to the relocation of operations, locating a firm's new production near other producers of similar products may be more beneficial as compared to locating this production near the firm's existing operations. One might envision that firms interested in entering integrated circuits would have heard about the advances occurring at pioneering integrated circuit firms and thought they might be able to gain an advantage by locating their integrated circuit operation in one of these clusters. This would have helped in terms of hiring away employees with key knowledge as well as with any informal exchange. In fact, some firms commented that their location outside of Silicon Valley made them feel disconnected from the industry and specifically the cutting-edge work that was occurring at firms like Fairchild (Saxenian, 1994, pp. 33-34).

Location decisions for manufacturing facilities require firms to make significant investments in facilities and labor which may be difficult to move to another location. As a result, it seems that, prior to entry into integrated circuits, firms would be likely to carefully analyze what the best location for manufacturing would be prior to first production, considering the baseline to be where the firm is already located. Assuming that the advantages of agglomeration outweigh costs related to relocation, we would expect to see producers from other regions relocate into agglomerated regions in order to realize these advantages.

In examining the location decision of firms, Alcacer and Chung (Alcacer & Chung, 2007, p. 774-775) provide some support for the idea that firms would locate their operations in regions where

they can enjoy advantages. While their study shows that firms will locate near sources of knowledge spillovers, it is limited to foreign firms that are considered first-time entrants into the United States. The authors note that the sample is limited to this group of firms because "prior investments can affect subsequent location choices and create dependence among observations by the same firm" (Alcacer & Chung, 2007, p. 766). Yet, if the potential knowledge spillovers were great enough, existing firms may decide to relocate in order to gain access to these spillovers. Given the nature of the transistor and integrated circuit eras of the semiconductor industry, we have a number of existing firms that we can examine to analyze whether they choose to relocate in order to take advantage of these knowledge spillovers.

One way to examine this is to look at all electronics (including components such as diodes, transistors, etc.) producers that locate their integrated circuit production in regions other than their previous production, focusing in particular on the role of semiconductor clusters. Clearly the firm has made a choice to locate its new operations in a new region. If these agglomeration advantages are evident, we would expect the vast majority of such relocations to be into agglomerated regions and that firms that "relocate" into these regions would enjoy advantages over other, non-relocated counterparts. We can examine this idea by testing the following hypothesis:

Hypothesis 9: Among integrated circuit producers with prior electronics production, firms that locate their integrated circuit production in semiconductor clusters, more specifically firms that relocate their operations to Boston, Los Angeles, New York and Silicon Valley, are more likely to out-perform other comparable firms, all else held constant.

While attributes such as location and experience have traditionally been included as part of an analysis of firm performance, the integrated circuit era is one that includes significant technological change, warranting some discussion and examination of how technologies pursued by firms influenced firm performance. Table 2 shows that, while there were a number of firms pursuing circuit miniaturization, many of the pioneering miniaturization firms focused on only one path to miniaturization. The different types of technology and competencies required to pursue each miniaturization path may have driven this need for focus. So, firms that didn't previously have broad experience in semiconductor technology or significant resources to devote to building the competencies required for multiple paths would find it difficult to pursue multiple miniaturization paths at once. Additionally, the resources required to pursue each path were most likely substantial, judging by the millions of dollars invested by the armed services for the development of modular and molecular circuits (Choi & Mody, 2009, pp. 16, 20). This may have prevented all except the largest firms from exploring multiple paths. Finally, the uncertainty of the miniaturized circuit market (Holbrook, 1999, p. 150) may have given firms caution in planning investments into miniaturization development. While the industry was certainly seen as one of growth, both the amount of growth as well as the time scale for this growth were uncertain. This uncertainty may have led firms to hedge their investment in miniaturization with investments in more traditional electronics. Regardless of the cause, it is notable that most pioneering firms were engaged in research and development in only one path to miniaturization.

This focused investment in miniaturization would not serve firms that chose losing miniaturization paths well in the future. The skills required for each type of miniaturization were

very specific to each approach, and the miniaturization path that eventually dominated the industry, monolithic integrated circuits (detailed below), was pursued by very few firms. As a result, only a select group of existing firms were able to leverage their miniaturization investments in the late 1960's when monolithic integrated circuits became the dominant technology. The firms that made the wrong bets lost valuable time in terms of development and competition and invested valuable resources in technologies that not only did not pan out, but also did not provide any real transferrable knowledge toward the approach that was eventually successful.

Until now, many of the predictions and hypotheses regarding the development of the semiconductor industry during the integrated circuit era are similar to those presented during the transistor era. However, a key difference between the transistor and integrated circuit eras is the gradual emergence of the monolithic integrated circuit, which would re-shape the electronics industry.

The following sections will examine the key players with respect to the development of the monolithic integrated circuit and the nature of the knowledge required to produce such devices. Additionally, how this knowledge may not have diffused, in the absence of an institution like Bell Labs in the transistor era, through standard spillover and collaboration mechanisms due to its tacit nature and the competitive nature of the industry will be explored. Finally, how spinoff firms may have helped to facilitate knowledge dissemination in a way that re-shaped the industry and led to the rise of a new industry cluster – Silicon Valley – will be discussed.

The Emergence of Monolithic Integrated Circuits

As described in the previous section, many different paths to miniaturization existed during the early years of the integrated circuit era. However, one path, monolithic integrated circuits, proved to be an overwhelming favorite of the industry over time, accounting for more than 90 percent of the value of shipped circuits by 1980 (Electronic Industries Association Marketing Services Department, 1980) . Given the impressive rise of this type of integrated circuit and its effective elimination of widespread interest in other approaches, this section will examine the developments that led to the dominance of this miniaturization approach during the integrated circuit era and its implications for the evolution of the industry.

As semiconductor technology evolved throughout the 1950's, systems were becoming more and more complex. While a variety of methods for producing electronic circuits, often referred to as miniaturization, were being pursued during the 1950's, two firms played key roles in the development of what was to become the dominant path: monolithic integrated circuits. These firms were Texas Instruments and Fairchild Semiconductor.

Integrated Circuit Development at Texas Instruments

The effort at Texas Instruments employed a great deal of knowledge from previous production, both within Texas Instruments and from Centralab, the former employer of a key TI employee, Jack Kilby. In 1958, Jack Kilby had been hired by Texas Instruments and assigned to work on TI's contract for an Army miniaturization program called Micro Module. However, Kilby was convinced that the Micro Module approach was not the appropriate path for miniaturization, in part due to his experience with a similar project at his former employer, Centralab, which was

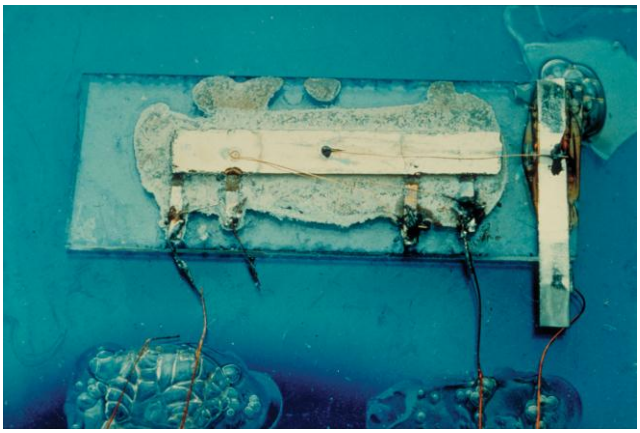
not successful (Reid, 2001, p. 75). As a new employee at Texas Instruments, Kilby had not yet accrued enough vacation to take the mandatory vacation period in the summer, leaving him as one of the few TI employees working over that period. He viewed this time alone at work as his opportunity to put aside the Micro Module project, noting that, "I felt it likely… that I would be put to work on a proposal for the Micro-Module program when vacation was over – unless I came up with a good idea very quickly (Reid, 2001, p. 76)."

During the vacation period, Kilby developed the idea of producing all of the components at once by integrating them within a single semiconductor substrate. The idea of producing the components simultaneously was also inspired by Kilby's time at Centralab, as the firm had been working on production methods that would place all circuit components on one ceramic substrate in a single manufacturing operation (Reid, 2001, p. 70). In theory, the same could be done with semiconductor materials. However, two components required for most electronic circuits, namely resistors and capacitors, had not yet been produced using semiconductors (Zygmont, 2003, p. 17). When Kilby presented his idea to his supervisor, Willis Adcock, he encountered significant pushback, with Adcock noting that the idea was "pretty damn cumbersome". To many, the idea seemed like a very expensive (and troublesome) way to produce circuits because resistors and capacitors produced using conventional materials performed well and were very cheap and reliable. Put simply, semiconductors were not seen as being the right material to produce good resistors and capacitors. While Adcock didn't necessarily believe that the idea would work, he was intrigued. After significant discussion between Adcock and Kilby, Adcock agreed to authorize the project if Kilby could make functioning resistors and capacitors out of silicon (Reid, 2001, p. 78).

Manufacturing resistors and capacitors out of semiconductor material wasn't a particularly easy task. To do so, Kilby worked with two technicians in TI's semiconductor laboratory. "Together they subdivided the surface of germanium bars that were about half the size of a stick of Dentyne. They scratched out and shaped regions of the material to function as capacitors, as resistors. (Zygmont, 2003, p. 20)" Once the devices were wired to the test circuit and performed appropriately, Adcock approved the overall project (Reid, 2001, p. 78), choosing a phase-shift oscillator as the demonstration circuit because it contained all basic active (diodes and transistors) and passive components (capacitors and resistors) (Reid, 2001, p. 79).

With the main uncertainty of the project eliminated, (diodes and transistors were routinely constructed with semiconductor materials at this point), Kilby's effort centered on constructing a circuit with all four components on the same semiconductor substrate, isolating the components so that they would each function properly, and connecting the components in the desired way on the surface of the substrate. Construction of the components was performed using mesa transistor techniques that had previously been employed at Texas Instruments (Lécuyer & Brock, 2010, p. 157) and were being used by others within the industry to reliably produce transistors. To isolate the components, Kilby relied on the practice that had been employed early on in demonstrating that capacitors and resistors could be constructed using semiconductor materials – he and his team etched away portions of the semiconductor to allow air to serve as the isolating agent. Additionally, connections between the components were achieved by wiring the components together on the surface of the substrate. Figure 10, below, is a photo of the original integrated circuit device produced by Kilby's team.

Figure 10: First Integrated Circuit, Designed by Jack Kilby at Texas Instruments.



Source: http://newscenter.ti.com/media/p/2884.aspx

The device wasn't particularly elegant, as Kilby noted in 2008:

> "I feel bad when I see it (the picture of the first integrated circuit)… it wouldn't have had to be that ugly. If I would have known that I would have to look at it for forty years I would have prettied it up a little bit". (Unknown, n.d., 2:31)

It did, however, function properly and served as the first demonstration of an integrated circuit device. Kilby's device was first witnessed by TI management on September 12, 1958, and quickly garnered interest from the military (Reid, 2001, pp. 79-80). Seeking to publicize the new potential product to prospective customers, TI disclosed the integrated circuit to the Air Force, the US Army Signal Corps and the Navy shortly after the first demonstration (Phipps & Laws, 2011, p. 10).

Overall, Kilby's invention moved the industry and the state of the art forward by including all of the required components for a circuit within a single semiconductor substrate, and TI is given

credit for inventing this portion of the technology required to produce today's integrated circuits. However, Kilby and his team relied on existing technology for both the isolation and connection aspects of the device. The biggest shortcoming of Kilby's earliest design was its failure to provide practical, easily producible interconnections, since the components "still needed to be interconnected by wires or by the equivalent of wires. Good interconnections were simply impossible because the early inadequately structured semiconductors at Kilby's disposal could not accommodate any integrated pathways." (Zygmont, 2003, pp. 22-23) The team at Fairchild Semiconductor would use their knowledge regarding the planar transistor to move the technology forward with regard to both isolation and connection and would become known within many circles as the inventors of the integrated circuit.

Integrated Circuit Development at Fairchild Semiconductor

While Texas Instruments' integrated circuit efforts focused on the integration of components within a single semiconductor substrate, Fairchild's development work focused on this objective as well as achieving mass-producible isolation and interconnection. It should be noted that, while Fairchild's developments will be presented as if they built upon the progress made at Texas Instruments, the two firms were working on integrated circuits independently of each other (Zygmont, 2003, p. 24).

The primary shortcoming of the monolithic integrated circuit efforts at Texas Instruments was the lack of an easy way to connect the various components of the integrated circuit for the circuit to function properly. Due to the etching used to isolate the components, each component had to be connected using wires or other physical connectors that had to be attached to the component.

While such an integrated circuit enabled circuit designers to produce smaller circuits with higher complexity, the tyranny of numbers problem still existed as physical connections had a limit as to how small they could be made. As a result, another method of constructing monolithic integrated circuits had to be developed in order to truly address the tyranny of numbers.

Fairchild's monolithic integrated circuit approach answered that challenge, and was closely based on the mesa and planar transistor work that had been performed at the firm in the late 1950's. Bob Noyce, one of the founders of Fairchild Semiconductor, had been privy to Jean Hoerni's planar transistor design and development work, and following a meeting with a potential customer, was inspired to apply much of Hoerni's work to a full integrated circuit (Lécuyer & Brock, 2010, p. 157). As a result, many of the lessons learned from and technologies developed for the planar transistor effort were applied to the planar integrated circuit development effort.

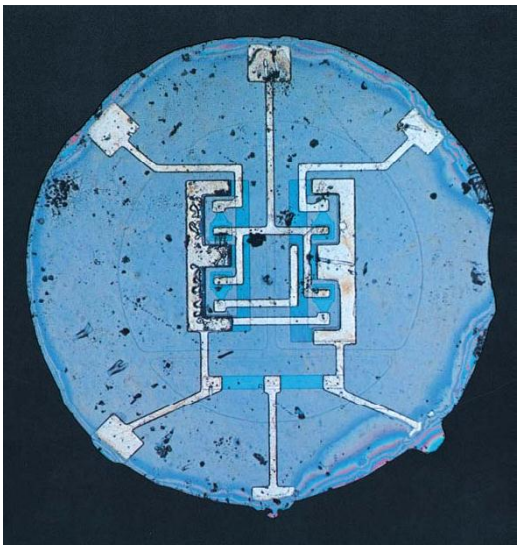
Planar transistors utilized silicon oxide for masking and diffusion operations in a way that was similar to mesa transistors. However, planar transistors broke from common practice by maintaining the silicon oxide layer on top of the substrate in order to protect the junctions in the transistor and increase reliability. Additionally, a metalized interconnect pattern was deposited on top of the silicon oxide to connect the various circuit components, eliminating the need for wire-based interconnections. This process borrowed aspects from the metallized connectors used in planar transistor production but used the same metallic materials to connect the various components in the circuit in a way that had not previously been done (Lécuyer & Brock, 2010, p.

157). It used the metallic interconnects as "wires of a sort" which could be mechanized and required incredibly small clearances, unlike traditional wires.

The use of the planar process in creating an integrated circuit that could be interconnected using silicon oxide and metallic deposition required further innovation at Fairchild. While Kilby at Texas Instruments had used mesa etching in order to isolate circuit components from each other, such an approach could not be used if the surface of the semiconductor substrate was to be maintained in order to create interconnections. As a result, Noyce and others at Fairchild had to develop another way to isolate the components within the circuit. Much as the individual components within the circuit were composed of junctions between p-based and n-based areas, Noyce created p-n junctions between the components, i.e., diodes, to eliminate current flow between components (Lécuyer & Brock, 2010, p. 157). This idea had been developed independently by Kurt Lehovec of Sprague Electric, but Lehovec had applied the idea to substrates in a way to create isolation chambers that would then house components created through the alloying process (Lécuyer & Brock, 2010, p. 157). Noyce's effort was particularly novel because it was applied to a planar integrated circuit, one that required isolation efforts such as the p-n junction between components in order to function properly.

Combining the methods used previously in planar transistor manufacturing to form components within semiconductor substrates and the new processes for isolation and interconnections, Noyce produced the first planar integrated circuit on May 26, 1960 ("Computer History Museum - The Silicon Engine | 1960 - First Planar Integrated Circuit is Fabricated," n.d.). Figure 11, below, is an image of the first planar integrated circuit.

Figure 11: First Planar Integrated Circuit, Designed by Jean Hoerni at Fairchild.



Source: http://dc265.4shared.com/doc/j_lGmPCn/preview.html

While TI's original design contributed to the technology by creating the first monolithic circuit – one with all of the components within a single semiconductor substrate – it was Fairchild's innovations that would propel the monolithic integrated circuit forward as a device with significantly greater reliability than other miniaturization concepts. Fairchild's contributions included protecting the fragile junctions of the integrated circuit by protecting them with silicon oxide, automating the connection of various components by overlaying a metallic layer on top of the silicon oxide, and isolating components within the circuit from each other.

These contributions addressed two very important challenges: the reliability of the device, which was significantly improved by protecting the junctions of the semiconductor device (similar to the reliability boost provided by planar transistors), and the manufacturability of the integrated circuit, which until that point had been made in a very labor-intensive manner. Because

components were formed by doping the substrate (as compared to alloying materials onto the substrate), and because the components were then connected using metals deposited using masks, many circuits could be made simultaneously using the same semiconductor wafer and a single manufacturing process. As a result, this manufacturing process could be mechanized in a fairly simple and efficient way as compared to the connections that had previously been made using wiring connections, sometimes applied by hand. Over time, these reliability and manufacturing advantages would propel monolithic integrated circuits into a leadership position within the industry, with these devices accounting for more than 90 percent of the value of all integrated circuits shipped by US manufacturers by 1980 (Electronic Industries Association Marketing Services Department, 1980).

Mismanagement, Disagreements, and the Role of Fairchild in Seeding a Region

As Fairchild developed into a repository of knowledge regarding monolithic integrated circuits, mismanagement plagued the firm, which resulted in a number of disagreements among Fairchild employees. The disagreements were numerous, and started in the early days of the integrated circuit, when Gordon Moore, the head of research at Fairchild was hesitant to commercialize the integrated circuit. Moore's inaction prompted two groups of top employees to form Amelco and Signetics in order to pursue integrated circuit commercialization (Lécuyer, 2006, p. 212). Commercialization intent was not the only source of disagreements, as complications regarding technology transfer between research and development, frustration regarding compensation and stock options with the east-coast parent of Fairchild, and the installment of new management

from competitor Motorola led to a number of spinoffs, including Intel, AMD and National Semiconductor, among others (Klepper & Thompson, 2010).[21] These disagreements, coupled with the increasingly tacit and complex nature of knowledge required to produce monolithic integrated circuits ultimately led to the formation of a number of spinoff firms in Silicon Valley, most of which became leading firms in the semiconductor industry.

Technologies Developed at Fairchild and Texas Instruments

The previous description of the invention and commercialization of monolithic integrated circuits does not entirely capture the complexity of the manufacturing process or the difficulties encountered by Fairchild and Texas Instruments. The production of a monolithic integrated circuit is very complex and involves a number of steps that in many cases had not yet been commercially used prior to the initial production by Fairchild and TI. In many cases, these steps or processes had been pioneered or invented by Bell Labs or other firms, but had not yet been used for production of complex and increasingly smaller integrated circuits. Such situations required Fairchild and TI engineers to improve the process for production and/or to adapt the process for semiconductors. The table below details all of the major processes used to produce a monolithic integrated circuit using the planar process, where the process originated, and additional work done at Fairchild or TI to adapt the process for use in producing monolithic integrated circuits.

[21] See Lécuyer (2006, pp. 258-264) for a more complete discussion.

Table 3: Major steps required for production of monolithic integrated circuits and sources of information for each step.

Area	Original Source	Additional Effort at Fairchild and/or TI
Boron and Phosphorous Diffusion	Bell Labs & RCA (Gaseous)	Fairchild: Scientific Lit for gaseous; Diffusion was used at Shockley; Engineered phosphorous solutions (powder and gaseous-diffusion) TI: Extensive experimental effort using multiple diffusants and both a single-step and dual-step diffusion process
Oxide Masking	Bell Labs	Fairchild: Access to Bell Labs Information from Shockley Experience TI: Oxide featured in original integrated circuit disclosures; further experimentation regarding oxidation thickness and resulting surface conditions
Photolithography	DOFL & Bell Labs	Fairchild: Shockley Experience Involved access to Bell Labs Information; Designed step and repeat camera
Photoresists	Eastman Kodak	Fairchild: Collaboration with Kodak to modify photoresists to work with silicon
Aluminum Contacts (Metallization)	Fairchild	Fairchild: Pioneered Aluminum Metallized Contacts TI: Further experimentation regarding contact deposition thickness, other contact materials and the use of back contacts
Isolation	Fairchild, Sprague, Bell Labs	Fairchild: Developed P-N Junction Isolation TI: Investigated Diffusion through an epitaxial layer, Triple Diffusion and Gold Diffusion for Isolation
Packaging & Assembly	Transistor Industry	Fairchild: Dual Inline Packaging TI: Flatpack Packaging
Testing Procedures	Transistor Industry	Both Fairchild & TI developed IC testing machines

Sources: Brower, Cragon, & Lathrop, 1965; "Computer History Museum - The Silicon Engine | 1965 - Package is the First to Accommodate System Design Considerations," n.d.; Lécuyer, 2006, p. 203; Lécuyer & Brock, 2010, pp.19-20; Millis, 2000, pp. 63-68

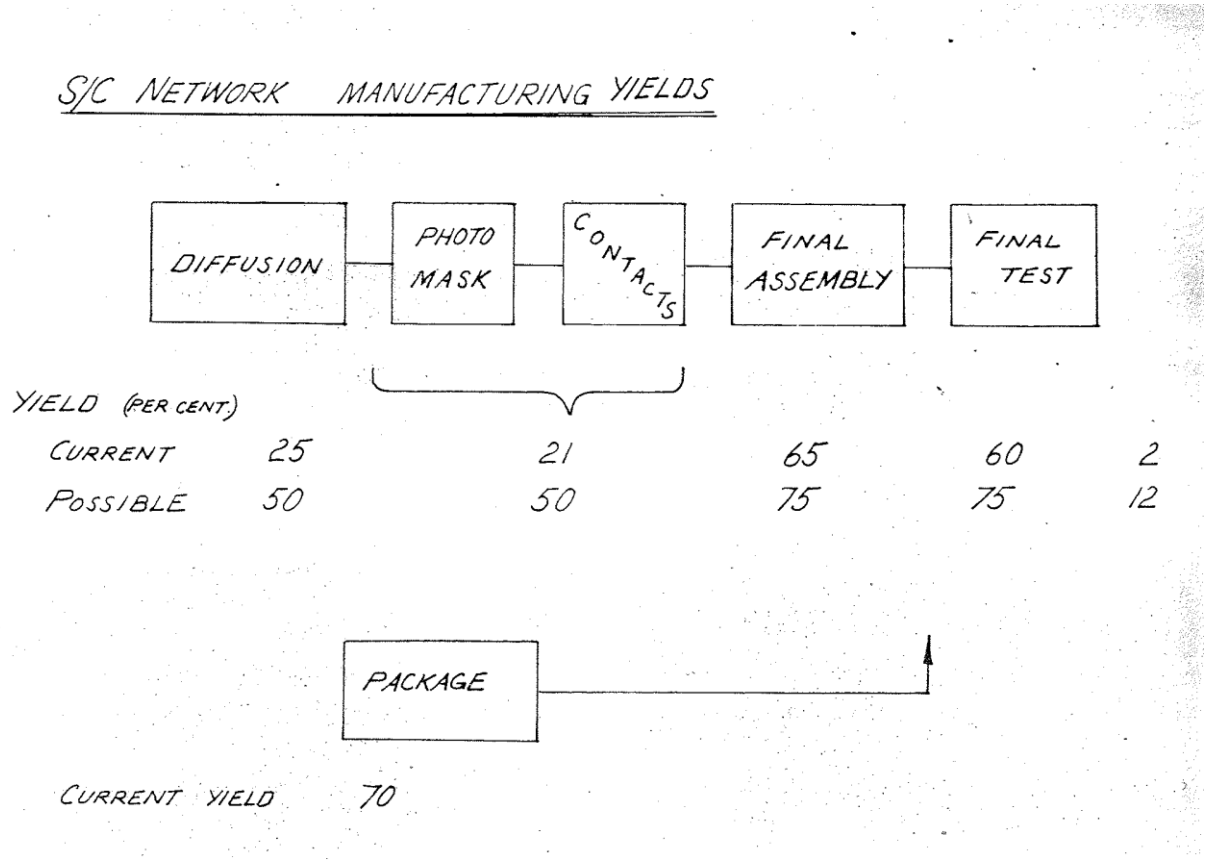
The development of these technologies and refinement of existing processes to allow for reliable production of integrated circuits was not without significant challenges. Many of the Fairchild developments noted above occurred while the firm was producing its first commercial production of planar transistors for Autonetics,[22] as previously detailed. Autonetics' contract for the Minuteman II rocket program would also play a key role in improving TI's manufacturing capabilities for monolithic integrated circuits. Initially, TI was unable to produce the quantity of devices required to meet the terms of the contract.

Correspondence between TI and Autonetics from September 10, 1963 references this production problem and indicates that only 10 of the 18 types of integrated circuits met or exceeded delivery goals set by Autonetics (Haggerty, 1963). It goes on to explain that the problems restricting production had not even been identified for 5 types of integrated circuits. The overriding production problem centered on manufacturing yields. As shown in the internal TI graphic below, the overall yield for integrated circuit production in the early 1960's was two percent

[22] Autonetics did not produce integrated circuits until 1970, following its acquisition by Rockwell. It is not clear that it used its experience with Fairchild and Texas Instruments to begin producing integrated circuits, since it only produced them following the acquisition. Additionally, several reasons may account for the firm's lack of integrated circuit production, including 1) a desire to maintain a strong supplier network for integrated circuits as the firm was a government contractor that required suppliers; 2) a strong desire within the industry for second-sourcing of key components, which would have included the integrated circuits used in missile systems for navigation; and 3) potentially an inability of the firm to work with silicon, which was the situation at American Bosch Arma in 1959 as documented by Lécuyer and Brock (2010, p. 157). Regardless, Autonetics was not producing integrated circuits until 1970 according to the *Electronics' Buyers Guide*.

(Unknown, n.d.-b),[23] similar to figures that have been mentioned for some of Fairchild's initial yields for planar devices.

Figure 12: Semiconductor Network Manufacturing Yields



Source: Texas Instruments records, DeGolyer Library, Southern Methodist University. (Unknown, n.d.-b)

In simple terms, for every 100 integrated circuits that were produced, only two were produced to specifications and performed satisfactorily when they underwent final testing. While a two

[23] The final number (two percent) can be found by multiplying all of the individual yields since the processes detailed are serial in nature.

percent yield is not unusual for initial production of new semiconductor technologies (Lécuyer, 2006, p. 152), such yields can create large problems for producers that have committed to producing substantial volumes of devices. Given TI's contractual production commitment to Autonetics, TI had two potential solutions: get Autonetics to decrease requirements to allow a larger portion of components to meet specifications or improve yields. With Autonetics' help and guidance, TI did both.

While TI had achieved short-term allowances to ship components to Autonetics under relaxed specifications, TI decided to investigate the root causes of the yield issues and to find a more long-term solution. It is not clear from the firm's documentation exactly what the impetus for the study was, as the primary report that documents the results of the study references an Air Force contract (Brower et al., 1965). However, it is clear that the efforts documented in this report most likely occurred in conjunction with the efforts to understand the yield problems. The research effort was very broad, with the report noting that the study examined process development for semiconductor manufacturing. However, it also notes that "process development work did not encompass all steps of processing but was restricted to those operations for which no suitable procedure existed" (Brower et al., 1965, p. 1). Even with this restriction, the study focused on many of the key processes required for planar integrated circuit manufacturing. The types of investigations for each process area as well as key results are indicated below (Brower et al., 1965):

- Lapping and Polishing: Mechanical and chemical polishing methods were examined. New chemical etch-polish methods were developed that significantly improved the output of the polishing process. Further development of mechanical polishing methods was discontinued.
- Diffusion: One-step and two-step diffusion processes were examined, along with various diffusion materials, temperatures and concentrations. The results of this study contributed to the science of diffusion, allowing TI and other firms to better predict the outcome of diffusion processes based on the operating conditions.
- Isolation: Four methods were examined (diffusion through the slice, diffusion through an epitaxial layer, triple diffusion and gold diffusion), and two methods were deemed to be practical (diffusion through an epitaxial layer and triple diffusion) with work on the other two methods being abandoned.
- Oxide Thickness: Various conditions that are typically encountered by the semiconductor substrate during diffusion were simulated in order to develop a better understanding of the various factors that affect the oxide thickness produced during diffusion.
- Contact and Lead Formation: Optimum conditions for contact and lead deposition were examined, including the design of filament used in the deposition process, the reproducibility of deposition between batches, as well as the spacing between the filament and the target surface for deposition.

Perhaps not captured completely in the description above is the extent of experimentation that occurred during this study.[24] The report features a number of curves that were derived based on experiments, as well as tables reporting many different test runs in order to help establish causality of modifying relevant characteristics. Experimentation was necessary because the science of planar integrated circuits had not yet been fully developed. The more complete understanding of planar manufacturing processes that was developed by TI through this study most certainly provided the firm with a competitive advantage.

[24] Similar experimentation seems to have occurred at Fairchild prior to and during production of planar transistors and early planar integrated circuits, as mentioned previously.

Importance and Adoption of Monolithic Integrated Circuits

Monolithic integrated circuits clearly required a great deal of technological innovation to make them practical and feasible, and the effort expended by Fairchild, TI and other firms would soon be rewarded within the industry. By 1980, monolithic integrated circuits would account for more than 90 percent of the value of all integrated circuit shipments within the United States (Electronic Industries Association Marketing Services Department, 1980), relegating the other types of integrated circuit technologies to niches within the industry or to history. Given the lucrative nature of monolithic integrated circuits, it seems likely that firms that were able to produce monolithic integrated circuits upon entry were more likely to out-perform other firms that were not producing products at the technological frontier upon entry and that integrated circuit manufacturers would be interested in adopting monolithic technology if possible to pursue the market advantages provided by operating at the technological frontier. The idea that operating at the technological frontier had distinct advantages can be tested by examining the following hypothesis:

Hypothesis 10: Firms that were producing monolithic integrated circuits upon entry were more likely to out-perform other firms that did not produce monolithic integrated circuits upon entry, all else held constant.

It seems from the activities of firms within the industry that production of monolithic integrated circuits was rewarded within the market, as many firms were pursuing monolithic integrated circuit production. Several integrated circuit firms were pursuing monolithic technology by licensing the technology from Fairchild and Texas Instruments, including Raytheon, RCA, IBM, Philips, Western Electric, and ITT (Texas Instruments, 1964). Other firms hired individuals with

integrated circuit knowledge in order to gain access to this body of knowledge. One example of this practice was the hiring of Frank Wanlass by both General Microelectronics and General Instruments following his employment at Fairchild to gain access to Metal Oxide Semiconductor monolithic knowledge (Bassett, 2002, pp. 149-156).

Assuming that the knowledge required to adopt monolithic integrated circuits can be facilitated by interactions between individuals and firms, we would again expect that firms co-located with other monolithic integrated circuit manufacturers would be more likely to learn earlier about this technology, and also to adopt it sooner. This expectation can be examined by testing the following hypothesis:

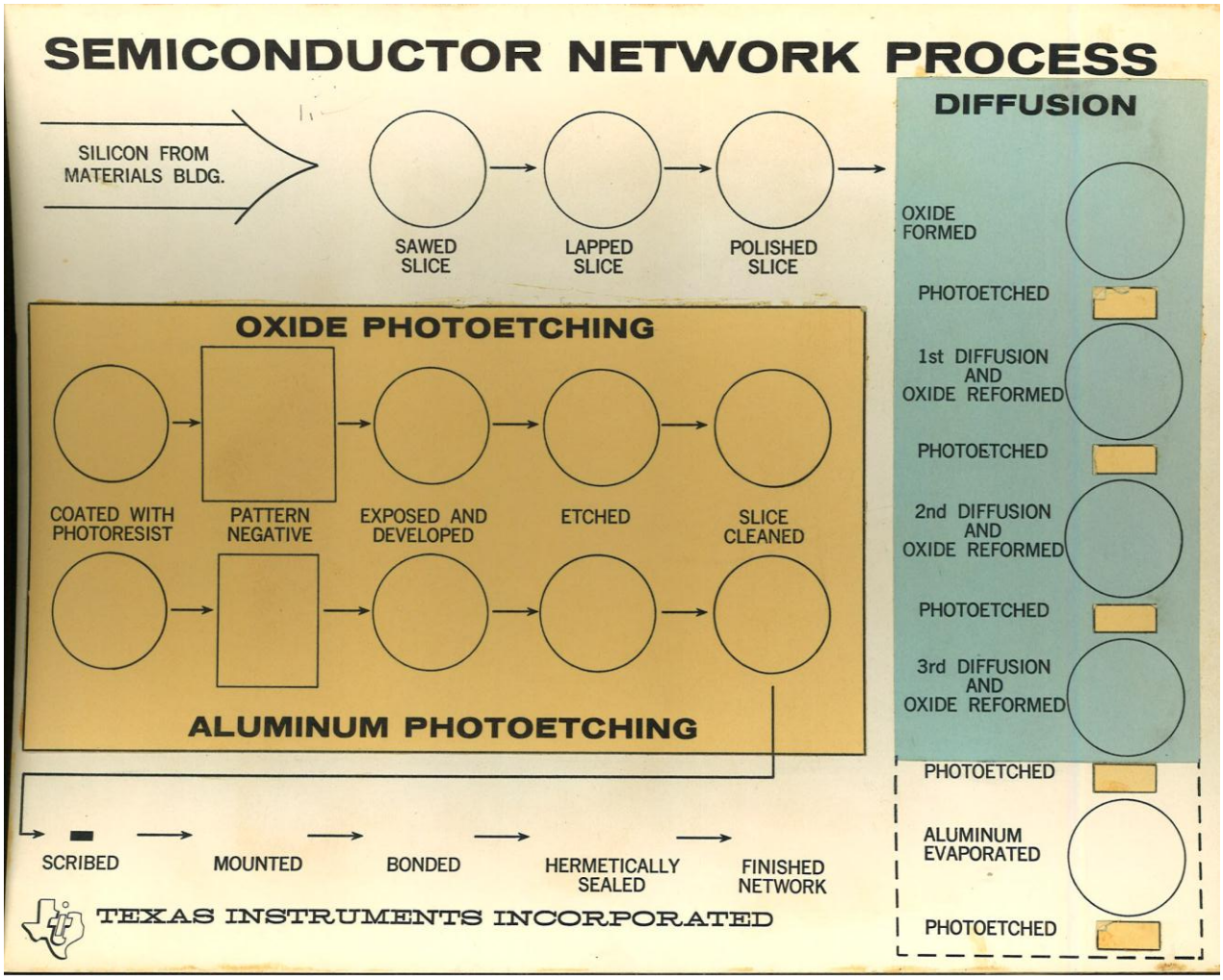
Hypothesis 11: The rate at which existing integrated circuit producers adopt monolithic ICs at any given moment is greater for producers located in clusters of monolithic integrated circuit producers compared to other similar firms located outside these clusters.

If findings support this hypothesis, it would provide evidence that a situation where the industry clusters are able to continue to strengthen as long as each cluster has access to cutting-edge knowledge would be feasible. Otherwise, firms located within a cluster without cutting-edge knowledge would be placed at a disadvantage to compete and develop in the future.

Increased Complexity

The fact that much of the technology for planar monolithic integrated circuits was developed within Fairchild and Texas Instruments through a great deal of experimentation leads one to question whether other firms had adequate access to the technology. In addition, the complexity of the knowledge and skills required for planar monolithic integrated circuit manufacturing must also be examined. Planar monolithic integrated circuit technology is seen as an extension of planar transistor technology, with some in the industry noting that "semiconductor microelectronics technology has grown out of the background of solid state research and manufacturing experience already in existence" and that "our most promising semiconductor integral circuits are being fabricated with the same materials and processes as those employed in making the epitaxial planar transistor" ("Address by Mr. James M. Bridges, Director, Office of Electronics, Office of the Director of Defense Research & Engineering, Before the Ft. Monmouth Chapter of the Armed Forces Communications Electronics Association," 1963). A process diagram detailing the manufacturing process of planar integrated circuits is shown below in Figure 13. For the most part, the same process diagram could be presented for the planar transistor production process, though the masking and diffusion would be of a simpler pattern and aluminum evaporation wouldn't be necessary.

Figure 13: Planar Production Process Diagram.



Source: Texas Instruments records, DeGolyer Library, Southern Methodist University. (Unknown, n.d.-c)

The similarities between other types of transistors and integrated circuits as well as planar integrated circuits, however, are far fewer. For both point contact and alloy junction transistors, none of the steps in the shaded boxes in Figure 13 are performed. The silicon (or other semiconductor material) still needs to be prepared by sawing, lapping and polishing to ensure the proper shape and state for processing, and following either the placement of the contacts or the

alloying of the materials to make the junction, the device would need to be properly mounted, sealed, and packaged, in a fashion similar to planar integrated circuits. However, the use of masks and photoresists, diffusion, oxides and photoetching are nowhere to be found in the processes used to produce point contact or alloy junction transistors ("Manufacture of Junction Transistors," n.d.).

Not only are the processes in the shaded regions not performed with other types of transistors, but the replacement processes are also different in two key ways. The replacement processes are physical processes and occur at the surface of the semiconductor substrate. For example, the leads are placed onto the semiconductor substrate in a fashion where they will form a junction that will allow the transistor effect to occur for point contact transistors. Similarly, the alloying material is placed onto the semiconductor substrate and then heated to allow alloying to occur for alloy junction transistors. Diffusion is an example of a process used in manufacturing planar integrated circuits which is chemical and occurs below the surface of the semiconductor. The dopant (labeled diffusant in the figure below) is heated to a temperature where it becomes gaseous, and then penetrates the surface of the semiconductor. Through the diffusion process, the molecular structure of a portion of the substrate is changed in such a way to create a junction between p-type and n-type sections. Figure 14, below, shows the equipment required for diffusion. Other processes used in the production of planar integrated circuits, such as photomasking, oxidation and etching, involve either the etching away at or adding to the surface, but none of these processes act at the surface of the substrate.

Figure 14: Equipment Used for Semiconductor Diffusion.

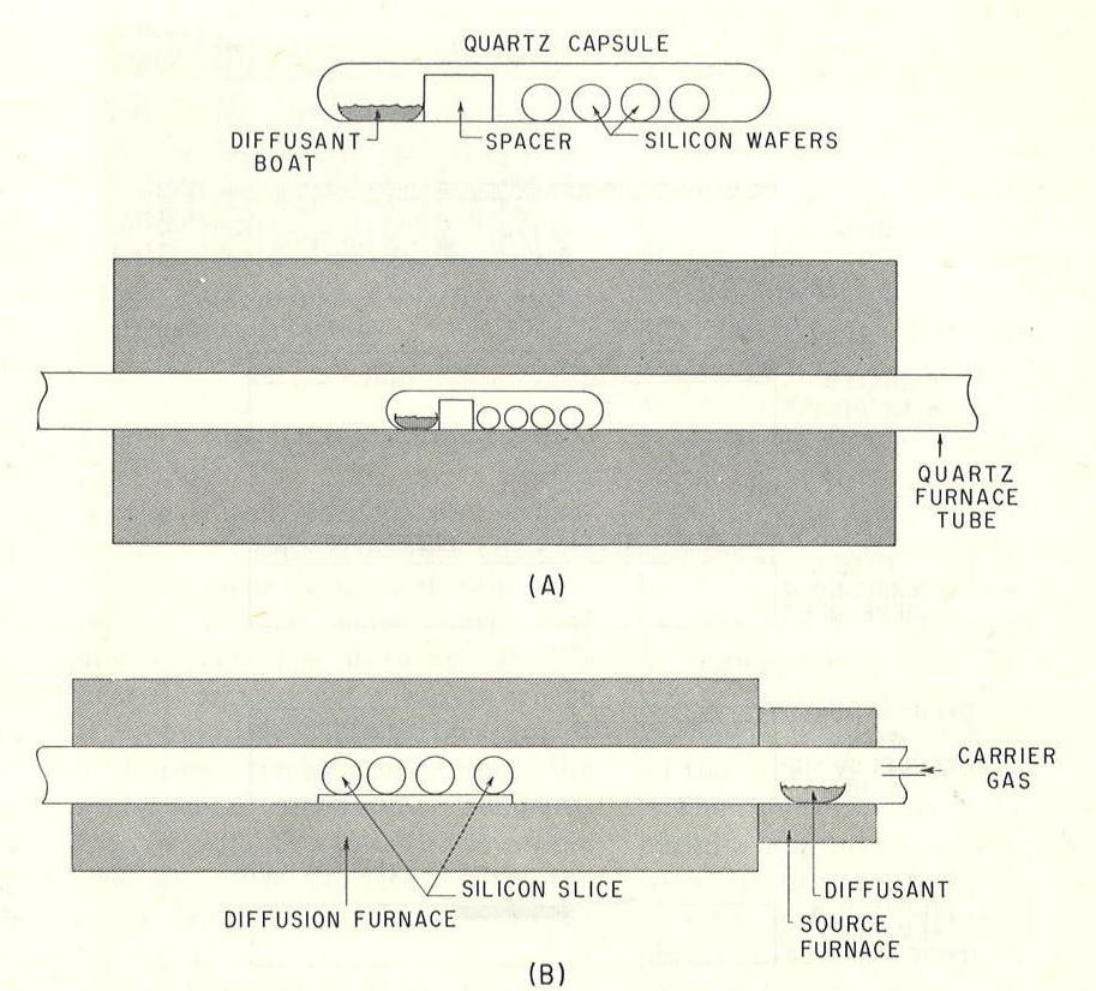


FIG. 4—Diffusion equipment showing (A) closed-tube system and (B) open-tube system

Source: Texas Instruments records, DeGolyer Library, Southern Methodist University. (Lathrop, Lee, & Phipps, 1960)

Given the similarities in processes used for planar transistors and planar integrated circuits, and the very different nature of these processes as compared to those used for other types of transistors, it seems plausible that firms that had previously produced planar transistors, such as Texas Instruments and Fairchild, would have an advantage when they attempted to produce

planar integrated circuits. While such related technical experience has been shown in other cases to positively affect a firm's propensity to enter other industries, it has also been shown to positively affect performance following entry into such industries (Klepper & Simons, 2000; Lane, 1989). Given detailed firm background data, this idea can be tested using the following hypothesis:

Hypothesis 12: Among integrated circuit entrants with previous experience, firms that had previous planar transistor production experience were more likely to perform better than other integrated circuit producers, all else held constant.

Further Semiconductor Innovations

Following the initial miniaturization work where the industry developed separate concepts regarding integrated circuits, future innovations were mostly centered on improving the monolithic integrated circuit concept and adapting it to different markets and applications. In many cases, these innovations occurred within firms that emanated from Fairchild Semiconductor through spinoffs.

MOS

One of the key technologies in advancing monolithic integrated circuits was metal oxide semiconductor (MOS) devices, initially realized as MOS transistors. This technology had wide-reaching ramifications for integrated circuits as MOS components were more easily produced using the planar process than other traditional semiconductor components. Additionally, the adoption of MOS components for integrated circuit manufacturing would lead to future innovations that would greatly increase the density of components on the integrated circuit, allowing the devices to be even more complex and powerful moving forward. As a result, the advent of the monolithic integrated circuit gave new life to a rather unpopular type of transistor invented in the late 1950's.

MOS transistors were very different from the more traditional bipolar junction transistors (BJTs), which were widely used in the industry at the time. The operation of MOS devices relied on a capacitor formed by the oxide coating of the semiconductor, sandwiched between a metal film and the semiconductor substrate (Bassett, 2002, p. 24), as opposed to relying on the junctions that were formed in the semiconductor substrate in the case of BJT devices. However, at the

time of their invention, MOS devices were slower and less stable than existing BJT devices, which discouraged many firms from pursuing MOS technology (Bassett, 2002, p. 13).

MOS devices had an advantage with respect to BJT devices, namely that they could be more easily constructed using the planar process, which was quickly gaining prominence within the integrated circuit field. The nature of their design allowed a large number of MOS devices to be integrated within a single circuit, which was a key metric of performance for integrated circuits. Additionally, the performance of MOS devices could be improved by shrinking their dimensions, providing an advantage for the technology as the interest in squeezing ever more components onto a single substrate increased over time. Given the numerous advantages, it seems that the decision to adopt MOS devices would have been a simple one. However, MOS devices relied on the surface of the silicon substrate in order to operate (Bassett, 2002, p. 31), an issue that had been cumbersome in the era of cat's whisker and early point contact transistors, as discussed earlier in this chapter. Additionally, while MOS production was a "subset of the steps used to produce bipolar transistors", transitioning from BJTs to MOS was difficult as "the fabrication processes required subtle but significant modifications to yield good MOS transistors" (Bassett, 2002, p. 7). Two firms led in the development of MOS devices: RCA and Fairchild. At RCA, this work was focused on MOS transistors as a potential "new fundamental building block" for ICs (Bassett, 2002, p. 44) as opposed to work on integrated circuits using MOS devices. This effort was pursued mostly at its Semiconductor and Materials Division in Somerville, NJ, and led by Karl Zaininger (Bassett, 2002, p. 42). Led by Frank Wanlass, work at Fairchild in the early 1960's aimed to control the surface in such a way that reliable MOS devices could be made.

Both firms were at least mildly successful in their pursuits of MOS transistor technology, allowing Fairchild and RCA to produce MOS transistor devices by 1963 (Bassett, 2002, p. 56).

Wanlass left Fairchild in late 1963, as he began to believe that the firm "was more interested in studying MOS structures than in selling MOS transistors" (Bassett, 2002, p. 110). This move would prove instrumental for two firms, General Microelectronics (GME) and General Instruments, as Wanlass' presence within those firms greatly influenced the firms' technological trajectory. GME was the first stop for Wanlass following his employment at Fairchild, where he continued to pursue the MOS technology to which he had greatly contributed at his previous employer. His pioneering work in the use of electron-beam evaporated aluminum to produce stable MOS transistors, and his realization of the role played by sodium in MOS reliability issues were at the cutting-edge of MOS research. Within less than a year GME had introduced a MOS transistor to the market, months ahead of his former employer (Bassett, 2002, pp. 150-151). By late 1964, Wanlass again grew impatient with his employer and decided to join General Instruments (GI), which was in the process of seeking semiconductor experts from multiple firms in a quest to enter integrated circuits (Bassett, 2002, pp. 152-154). Within a few years Wanlass and his colleagues at GI were able to introduce new MOS integrated circuit products, including larger shift-registers and a digital differential analyzer, both of which were advancements to the state of the art in terms of MOS technology. Once again, in 1970, Wanlass decided to move on, leaving GI.

The importance of Wanlass in the dissemination of MOS knowledge is hard to overstate, as Wanlass trained many engineers within Fairchild, GME and GI, and imparted his knowledge on

many others located in other firms (Bassett, 2002, pp. 155). Given the "black magic" nature of MOS fabrication, which as one engineer maintained, "rested on a knife edge of technique that was liable to sudden and inexplicable upset" (Bassett, 2002, pp. 113), there was a great deal of tacit knowledge involved in the process. Much of this knowledge could only be transmitted in person (Bassett, 2002, pp. 155), allowing for the buildup of MOS knowledge within the firms where Wanlass had worked. Given the leading nature of Wanlass' work, this knowledge was highly coveted by others in the industry. Although Wanlass may not have been as well-known as some others in the industry, firms like Intel attempted to hire him to gain access to his store of knowledge (Bassett, 2002, pp. 156). Ultimately, Wanlass' employment resulted in the buildup of complex, tacit knowledge within Fairchild and some of the "Fairchildren", including Intel following a mass exodus from Fairchild (Bassett, 2002, pp. 173). This left many existing firms at a disadvantage with respect to a technology that would become increasingly important with the further development of monolithic integrated circuits.

CMOS

Monolithic integrated circuits were not able to be used for low-power applications as the devices required power at all times that the circuit was operating. By changing the design of MOS devices, however, a new type of semiconductor device was invented that would allow for circuits that required less power, opening up new markets to integrated circuit devices.

Two types of MOS designs existed: p-channel components were simpler to design, while n-channel components switched faster. Typically only p-channel or n-channel components were used in a single device (Bassett, 2002, p. 163). However, by matching complementary channel

components (i.e. matching p-channel and n-channel components together), the power requirements of MOS circuits could be greatly reduced, with some estimating that the reduction could be six orders of magnitude (Bassett, 2002, p. 162). This use of complementary channel components within a single MOS device was initially disclosed by Frank Wanlass in 1963, while he was still at Fairchild (Bassett, 2002, p. 162), and was called complementary metal oxide semiconductor (CMOS).

The main advantage of CMOS over other designs was the low power requirement, which occurred because CMOS chips only needed to be powered when transistors were in the process of switching between states. This is unlike other types of transistors such as bipolar junction transistors (BJTs) which require power whenever the device is in the "on" state. This power advantage allowed solid state devices to be applied to many applications where they previously could not perform adequately, including electronic wrist watches.

While CMOS was invented at and disclosed by Fairchild, the firm was looking for devices that could be produced in the short term. It thus overlooked CMOS for other alternatives. Another firm, RCA, embraced CMOS and produced its first CMOS commercial circuits in 1967. However, RCA's actions were "right on the wrong time scale", with CMOS languishing until it would become the only viable MOS technology in the late 1980's following a long period of exploration by a number of firms (Bassett, 2002, p. 165).

Memory

In 1968, two of the founders of Fairchild, Gordon Moore and Robert Noyce, frustrated with issues there, set out to create a new company operating in a product area they believed would be promising. This new firm, Intel, would pursue semiconductor memories, beating to market many of the other established firms that had been pursuing such memories. While many of the underlying technologies required to produce semiconductor memories existed, one key difficulty remained when Intel was founded: the alignment of the transistors to achieve the density that would allow for economically feasible semiconductor memory devices. Building on work that had been done at Fairchild, Intel overcame this difficulty to create the first commercial semiconductor memory.

While a number of solutions to transistor alignment were available, Fairchild began pursuing one avenue, silicon gate MOS, in late 1967. Instead of using aluminum to create the gate electrodes, silicon gate MOS technology used silicon for this purpose, providing self-alignment of the MOS transistor terminals, which resulted in very tight spacing between transistors. This advancement, as well as reliability gains realized because the silicon gate material could be applied beneath the oxide layer, would eventually make memory production economically viable (Bassett, 2002, p. 177). When Noyce and Moore left Fairchild to form Intel, however, a number of problems remained with the technology. Recruiting a number of Fairchild personnel to join them and structuring the firm so that research and development were done within manufacturing departments, Intel personnel worked aggressively through these issues. By September 1969, Intel had "developed a silicon gate process with good yield and reliability" (Bassett, 2002, p. 189). Following Intel's introduction of its first memory products, the firm became a leader in

semiconductor memory products, housing a great deal of knowledge regarding the silicon gate MOS process within the firm.

Microprocessor

Started as a memory company, Intel would soon transition to producing a much more complex device, the microprocessor. The firm pursued the creation of standardized semiconductor products, or products that did not involve customization from customers looking for special circuits and devices. This was an approach they adopted from their days at Fairchild (Bassett, 2002, p. 263), which led them to semiconductor memories. But future interactions with customers would soon convince Intel's founders that other applications of standardized devices presented significant market opportunities. One such customer interaction, with the Japanese calculator firm Busicom, resulted in the invention of the microprocessor.

Busicom was interested in ordering 7 different chips that would be used in a family of new calculators. With Intel's primary MOS designers occupied with memory design work, another employee, Craig Hoff, worked with Busicom to develop an appropriate design to meet their needs. After analyzing Busicom's needs, Hoff recommended an approach that standardized the chip that would process the inputs to generate outputs and use separate memory chips to create a standardized "brain" for the calculator family that could be programmed. Busicom decided to proceed with this approach (Bassett, 2002, p. 268). The development built on Intel's previous work with silicon gate MOS technology, which made "large, complex chip[s] like the 4004 or the 8008 [early microprocessors] manufacturable" (Bassett, 2002, p. 269) and employed the knowledge of Federico Faggin, who performed the initial development work on the silicon gate

MOS technology at Fairchild (Bassett, 2002, p. 268). As a result of this effort, Intel was finally able to produce a microprocessor that met Busicom's needs and started a revolution of sorts.

Intel's first "computer on a chip" was introduced in late 1971, and Intel's dominance of this space continues even today. Other startups attempted to produce microprocessors in the late 1960's, including Viatron and Four-Phase Systems (Bassett, 2002, p. 253), but neither firm was successful in launching competitive products to Intel's microprocessor. Today, despite formidable competitors such as Motorola and AMD (another Fairchild spinoff), Intel remains the market leader in microprocessor technology.

Each of these product categories – MOS, CMOS, Memory and Microprocessor – represent key technologies and devices that were built on monolithic integrated circuits and would account for a large portion of sales of integrated circuit devices. Furthermore, the designs for these devices were, in many cases, reliant on both new innovations as well as predecessor technology (e.g., CMOS was reliant on MOS). As a result, each design required that firms possessed the requisite knowledge that drove previous innovations, which made it difficult for firms to catch up to the state of the art if they had not produced pioneering devices in the past. In part due to this reliance on cutting-edge knowledge in moving the technology forward, Fairchild – and its spinoffs – played a key role in the development of each of these technologies and devices, as is evident in the description of the technological developments. While many diversifying and existing electronics firms were interested in pursuing integrated circuits, many of them (identified in Table 2) made poor choices regarding their initial efforts, which limited their

ability to make significant advancements in monolithic integrated circuits and products that built on this technology. These firms were at a disadvantage once monolithic integrated circuits began to dominate the semiconductor industry.

The Role of Spinoffs in Integrated Circuit Production

Throughout this chapter, most of the discussion regarding firms and the previous production experience of firms has been focused on existing firms that were diversifying into either transistors or integrated circuits. However, one of the two leading firms in integrated circuit development, Fairchild Semiconductor, was not an existing firm. Rather, Fairchild was a very different type of firm – a spinoff firm.

As described in the previous chapter, spinoff firms are firms formed by individuals previously employed within the industry (Klepper & Thompson, 2010, p. 2). Heritage theory maintains that spinoff firms will be influenced by the employment history of the firm founder(s), both in terms of production as well as performance. Additionally, heritage theory predicts that agglomerations can occur around early, highly-successful entrants as the spinoff process causes individuals from the early, highly-successful entrant to form spinoffs nearby that will also perform well. The theory suggests that spinoff firms will enjoy advantages over other new entrants, and in some cases existing firms, in several areas, including access to the superior technological and market knowledge existing in their top performing parents. This section will examine the specifics of the integrated circuit era of semiconductor electronics in order to form hypotheses using both heritage theory and the context of the industry.

Technology Knowledge

In industries where technology is a key component of the business, knowledge regarding the technology can be particularly important in starting a new venture. This knowledge could be associated with how the technology works, how products using the technology can be

manufactured, or advantages or limitations of the technology that are otherwise not obvious. Unlike the transistor era, there was not the same effort to broadly disseminate technological information throughout the industry. One reason for this culture of secrecy was that there was no impending anti-trust action to entice firms to share "everything they knew" about integrated circuits. Additionally, two firms in particular had invested large amounts of resources into developing integrated circuits and were interested in appropriating the rewards from their investments. While the average diversifying firm or new firm did not have access to cutting-edge information in the integrated circuit era as it may have had in the transistor era, workers in the firms leading the technology development did have access to critical knowledge. A spinoff firm created by one of these employees would be able to take advantage of technology that had been applied or developed at the parent firm of the founders. Two examples of technology knowledge being transferred to spinoffs through founders include Fairchild and Rheem Semiconductor, a spinoff of Fairchild.

In the case of Fairchild, the founders were able to use their experience in manufacturing with silicon at Shockley to produce silicon transistors in their new venture. Specific tasks that used technological knowledge from Shockley included the creation of a clean room environment and building both silicon growing and diffusion furnaces (Moore, 1998, p. 54). While these tasks may not seem extraordinary at this point, it is important to note that at the time Fairchild was founded this information was not widely disseminated, making the team's previous experience a valuable commodity.

Shortly after Fairchild introduced the silicon mesa transistor, the firm would be affected by a defection that would change its strategic direction: the vice president and general manager Ewart Baldwin decided to leave and form Rheem Semiconductor (Lécuyer & Brock, 2010, p. 162). Those at Fairchild were concerned that Baldwin would use his knowledge of the silicon mesa transistor to compete directly with Fairchild, as he indicated that he would (Lécuyer & Brock, 2010, p. 162). And they had good reasons to be concerned as Baldwin took with him his direct experience at Fairchild, as well as a copy of the detailed Fairchild process manual, which was later discovered at Rheem (Lécuyer & Brock, 2010, p. 162). The threat of a direct competitor using knowledge developed at Fairchild led the remaining staff to fully embrace the planar transistor because Baldwin and his colleagues at Rheem had left shortly before the disclosure of this invention (Lécuyer & Brock, 2010, p. 162). In both Fairchild and Rheem spinoff cases, the founders were able to take valuable technical knowledge with them and used this knowledge to pursue business opportunities within new ventures.

Market Knowledge

While knowledge about the technology may be very valuable to engineers, knowledge about the market where the technology can be best utilized is particularly valuable when a new market is being established, where many questions about the technology application subsist. This market knowledge can include information such as user needs, sales and distribution channels, and supplier connections, and can provide firms with an important advantage with respect to matching the technology to the right market in the right way.

While firms diversifying from electronics production to transistor production possessed a number of complementary assets (sales channels, marketing, etc.) due to the one-to-one replacement nature of transistors for vacuum tubes, the same was not true for transistor firms looking to diversify into integrated circuit production. Instead, the integrated circuit era of semiconductors involved the creation of an entirely new market. Where firms had traditionally sold components to original equipment manufacturers that would then design circuits specific for their applications, integrated circuit production required firms to design and produce integrated circuits to be sold to original equipment manufacturers. As a result, integrated circuit manufacturers used sales channels that were different than component manufacturers, had to communicate very differently to customers regarding device specifications, and ultimately were taking away jobs from their previous customers (Kowalski, 2011b). In doing so, IC makers captured a bigger piece of the value chain because the IC incorporated circuit design that was previously done by the electronics firms. As mentioned previously, Tripsas' analysis (1997) showed that diversifying firms were only able to effectively survive technological change when their complementary assets were still valuable following the technological change. The transition to integrated circuits is a case where the complementary assets of transistor (and other component) firms were no longer valuable. This created an opportunity for new firms to unseat diversifying firms. Spinoff firms created by founders with the right combination of skills and experience in the parent firm, while falling into the new firm category, enjoyed an advantage in that they had already been exposed to the integrated circuit market. This provided spinoffs with valuable knowledge about the market and complementary assets that were applicable and valuable to the integrated circuit market.

One example of a spinoff that was able to adapt the technology to the market using information about the customer is Signetics, a Fairchild spinoff. While Fairchild today is known as the firm that invented the integrated circuit, shortly after the invention of the device there was some resistance from Fairchild management to pursue the commercialization of the device (Lécuyer, 2006, p. 212). A group of three key employees from Fairchild involved in the development of the IC founded Signetics to produce integrated circuits. Since there wasn't widespread demand for such devices at that time, the team employed knowledge of the users and worked to develop integrated circuits for logic applications that used the diode transistor logic configuration, one that was very familiar to many of the prospective users of the device (Lécuyer, 2006, p. 212). Had the team not been exposed to systems firms as customers during their experience at Fairchild, it is possible that Signetics would not have been able to adjust in such a way as to rescue the firm and become a leading producer of integrated circuits (Lécuyer, 2006, p. 212).

Management Knowledge

Finally, both technical and market knowledge is not beneficial to the spinoff firm if the new firm cannot be managed and led appropriately. While Shockley Semiconductor was founded by a co-inventor of the transistor, had access to substantial knowledge from Bell Labs, and employed some of the most well trained scientists and engineers, this knowledge and talent was not enough to guarantee the firm's success. Rather, Shockley's bizarre management practices and disagreements with employees over technology strategy drove the "traitorous eight" that founded Fairchild out of the company (Lécuyer, 2006, pp. 135-138), and Shockley's firm would not recover following this mass exodus (Shurkin, 2006 pp. 185-187). Novel efforts at Fairchild aimed at building markets for new devices, such as application notes, were used at Fairchild

spinoffs (Lécuyer, 2006, p. 271). Additionally, Gordon Moore, widely known for his invention of Moore's Law and leading Intel (another Fairchild spinoff), noted that he and the other Fairchild founders learned how to be technologist-managers, a skill that they were able to utilize after leaving Fairchild to found other spinoffs (Moore & Davis, 2004, p. 13).

Responding to a lack of knowledge about silicon transistors in the market, Fairchild became a market leader in application engineering practices. Put simply, these efforts worked to communicate electrical characteristics and potential uses of components to engineers in the field given the novel nature of Fairchild's products. One example of this communication was the production of application notes, which "were meant to show commercial customers how their problems could be solved better with silicon devices" than with conventional components (Lécuyer, 2006, p. 197). The applications notes included information on how to use Fairchild's devices, even including sample circuit designs to assist engineers in using the devices, and were successful in helping to sell Fairchild's products (Lécuyer, 2006, p. 197). After Fairchild employees founded spinoff companies, many emulated this practice in an effort to increase their sales, with National Semiconductor serving as an example of such emulation. Following the practice that had been established at Fairchild, employees at National Semiconductor published application notes which described the main characteristics of circuits in order to increase the usability of these devices to engineers. Lécuyer notes that these notes, as well as other application engineering efforts at National were "a major factor in opening up the industrial market for microcircuits" (Lécuyer, 2006, p. 271).

Technologist-managers were trained scientists and engineers who were also able to lead others within organizations, and Moore specifies three key skills that technologist-managers needed to possess: managing personnel and the firm, structuring a technology business, and managing discovery. Moore admits that these skills hadn't yet been taught widely to technical employees in the late 1950's as Fairchild was being launched, and indicates that it was this skillset, developed at Fairchild and taught to many leaders within the firm, that allowed many of the "Fairchildren" spinoffs to thrive following their founding. Moore indicates that these experiences at Fairchild allowed his spinoff, Intel, as well as many other Fairchild spinoffs, to succeed. In fact, this effect was so widespread among Fairchild spinoffs that many within the industry referred to the firm as Fairchild University and spoke of its importance as an educational and managerial training ground (Moore & Davis, 2004, p. 14).

While the context explained above illustrates well how founders of spinoff firms are able to take knowledge regarding technology, markets and management from their employers, the question remains as to what consequence this knowledge transfer has. Broadly, how are spinoff firms able to take advantage of this knowledge, and, aside from the examples cited above, does this knowledge influence the strategy and performance of spinoff firms? With the appropriate information regarding the background of firms, production records and performance data, we should be able to examine what effects, if any, the spinoff process has on the firms that are formed through that process.

It seems that there are specific situations that might be more conducive to spinoff firms having a more pronounced advantage as compared to other new entrants. These situations would include

those where it is difficult or overly costly to gain access to knowledge that is needed to compete within the industry. These are found in industries where:

- There is significant secrecy regarding technological knowledge and licenses to such knowledge are heavily controlled and/or confined to specific partners.
- Technology is developing rapidly and thus rendering existing technological knowledge that might be disseminated through reverse-engineering and other mechanisms obsolete.
- The knowledge required to compete is tacit, and thus can only be conveyed through experience.

If any of these situations occurred within an industry, causing knowledge to be difficult or overly costly for other firms to access, spinoff firms may serve as a key conduit of knowledge within the industry since spinoffs allow firm founders to overcome many of these situations.

The integrated circuit era of the semiconductor industry presents us with a particularly interesting case within which we can examine how spinoff firms serve as a conduit for knowledge transfer within an industry. We have seen that a significant portion of technologies leading up to the planar transistor and monolithic integrated circuit were developed within two firms – Fairchild and Texas Instruments. The additional complexity of skills and knowledge required to produce monolithic integrated circuits, as well as the concentration of these skills and knowledge within two firms seems to have created a situation where a large portion of the necessary skills and knowledge required to successfully produce integrated circuits was tacit and required substantial connections to these two leading firms. If this is the case, spinoffs would have served as an important conduit for knowledge transfer within the industry during the integrated circuit era.

While the transfer of the three types of knowledge (technology, markets and management) may not be observable jointly, it seems that similarity in technology and market knowledge between two firms should be observable through the types of products that are produced by the firms. Clearly technology knowledge would need to be of similar levels, and, likewise, if the products have specific market applications, market knowledge would also be somewhat similar between the parent and the spinoff firms. As a result, the similarity in technology and market knowledge between two firms can be examined, or put differently, the extent to which technology and market knowledge is indeed transferred between a parent and its spinoff can be examined. This can be done by examining similarities between the product portfolios of the two firms.

It should be noted that the founder's knowledge transferred through the spinoff process is brought with the founder when the firm is first established. As a result, evidence of such knowledge should be evident at the time of the firm's founding or within a few years. Monolithic integrated circuits were an extremely important and valuable type of product to semiconductor firms, and therefore we should expect that information about monolithic integrated circuits would be some of the most likely to be disseminated from the parent firm when spinoffs are formed. Additionally, it has been suggested that the increasing complexity of the technology coupled with secrecy around the technology may have hindered the effectiveness of other mechanisms to transfer knowledge. Specifically, given that a spinoff is more likely to produce the same type of product as its parent firm (Anton & Yao, 1995; Cassiman & Ueda, 2006; Franco & Filson, 2006; Klepper & Sleeper, 2005; Klepper & Thompson, 2010), how the spinoff process allowed for knowledge dissemination in this broad product category can be tested using the following hypothesis:

Hypothesis 13: Firms founded by individuals coming from a parent firm with previous experience producing monolithic ICs are more likely to produce monolithic ICs than firms with other backgrounds, all else held constant.

This hypothesis could also be refined to test the knowledge dissemination between parents and spinoffs for all types of integrated circuits, which would allow for varying effects of spinoffs on dissemination depending on the characteristics of the various technologies and markets. A hypothesis to be tested in examining this would be as follows:

Hypothesis 14: Among spinoff firms, the probability that a spinoff firms produces a given product will be greater if the spinoff's parent previously produced the product, all else held constant.

Silicon Valley in particular had a disproportionately large number of spinoffs, which attracted the attention of Don Hoefler, an electronics journalist who first used the term "Silicon Valley" in print to describe the region. In a series of articles about the region, Hoefler identified a number of spinoffs in what would serve as the first instantiation of the "Silicon Valley Genealogy" (Hoefler, 1971a, 1971b, 1971c). This genealogy would later be developed by the industry group SEMI into a visualization of a great number of spinoffs that were formed within the region. If support is found for the previous two hypotheses, it draws attention to the large number of spinoffs that existed in Silicon Valley and the role that they may have played in disseminating technological knowledge. Given the large number of spinoffs that existed within Silicon Valley and Fairchild's location there, it seems reasonable to expect to see a disproportionate amount of firms in the Valley that produced monolithic integrated circuits upon entry, many of which are likely to be spinoffs. To examine how spinoffs may have facilitated this knowledge dissemination in Silicon Valley, we would test the following hypothesis:

Hypothesis 15: Silicon Valley entrants in the IC industry are more likely to produce monolithic ICs at entry than entrants in the other clusters and elsewhere.

The previous hypotheses examine the transfer of technology and market knowledge between parent and spinoff. However, if knowledge regarding the management of the firm is something that individuals learn during their experience at the parent firm and then take with the individual to the spinoff firm upon founding, then we would expect to see that the performance of the parent firm would have some predictive ability with respect to the performance of the spinoff firm. While certainly not every spinoff from a high quality parent firm will be a high quality spinoff (Klepper & Thompson, 2010), the spinoffs from a high quality parent would be more likely to be exposed to high quality managerial processes and thus should be more likely to utilize this knowledge than other firms, all else held constant. In order to examine the dissemination of management knowledge through the spinoff process, the following hypothesis could be tested:

Hypothesis 16: Among integrated circuit producers, spinoffs that have a high quality parent firm are likely to out-perform other firms, all else held constant.

Unlike the transistor era, it appears that the nature of the technological change, in terms of its increasing complexity, the market that was being created, and the location of key technological knowledge, namely within Fairchild and other pioneering firms, allowed Silicon Valley to emerge as the leading region for integrated circuit manufacturers. This emergence, however, may have been driven by the learning and dissemination process facilitated by spinoff firms, as opposed to through agglomeration economy effects traditionally thought to have been vital to the region's existence. The next several chapters will examine these hypotheses to try to determine

to what extent each theory correctly describes the development of the semiconductor industry during both the transistor and integrated circuit eras.

Stepping Back

While Chapter 2 presented existing economic theories that describe firm entry and performance, these theories cannot blindly be applied to any given industry or technology. Throughout the chapter, two eras in the semiconductor industry have been examined and existing theories on firm entry and performance have been combined with the technological stories and industry conditions of the transistor and integrated circuit eras to generate specific predictions and expectations.

Two key themes have emerged from this exercise. First, it has been found that specific conditions at the intersection of technology, strategy and public policy may strengthen or mitigate the effects of existing theories. It is believed that the openness of Bell Labs regarding transistor technology may have served as a mechanism to "level the playing field" between firms interested in producing transistors. This openness, which allowed firms across the country to access Bell's knowledge, was influenced by an anti-trust case launched by the US government. Additionally, as is the case for firms that do not "operate in so wide a field of economic activity that they are able themselves to benefit directly from all the new technological possibilities" as a result of their research (Nelson, 1959, p. 302), Bell's openness was motivated by the realization that it alone could not fully realize and exploit all possible transistor technologies. While firms located near Bell Labs in the New York area may have still benefitted from knowledge spillovers through mechanisms such as employee mobility, Bell's practices are expected to have made co-

location a second-order effect. Additionally, transistor spinoffs that emanated from Bell Labs had similar access to critical Bell transistor knowledge as other licensees due to its open and forthcoming exchange of such information. The technological knowledge gained through previous employment that normally might have been considered a competitive advantage was available to a large number of firms, placing spinoff firms at an equal footing as other firms with access to Bell Labs' transistor technology and at a disadvantage with respect to the complementary assets that many electronics diversifiers possessed.

While the specifics of the industry and technology may have reduced the importance of agglomeration effects in the transistor era, it appears that this was also the case during the integrated circuit era, though driven by different factors. As the industry evolved and monolithic integrated circuits gained preeminence, the nature of knowledge in the IC era that emerged in the closed environments of Fairchild and TI limited spillovers, strengthened the effects of heritage, and provided close to ideal conditions for the emergence of spinoffs. The symposia and site visits of the transistor era did not occur in the IC era, and the increasingly tacit nature of the knowledge required to produce integrated circuits required prior experience in developing and producing ICs to gain access to this knowledge. Given the lack of openness at an industry level, direct experience working at another integrated circuit firm emerged as the primary way to acquire sufficient knowledge about the technology and the industry. This, coupled with the fact that integrated circuits were a very lucrative, quickly developing product area, fueled disagreements among individuals at firms, providing the impetus for many to start their own spinoff firms. While agglomeration effects may still be present for firms that collaborate within

regions, those effects are expected to be dwarfed by the effect of the spinoff mechanism as a way to disseminate crucial knowledge within the industry.

The specific expectations and predictions that emerge from this analysis yield the second theme: the era associated with the development of the transistor is very different from the era of the integrated circuit. Even though the advantages of being located near Bell may not have been realized during the transistor era, it is still expected that firms with previous electronics experience, at this time clustered in Boston, Los Angeles and New York City, would be most likely to begin producing transistors and succeed following entry. This prediction would yield the build-up of transistor firms within the existing electronics clusters, further developing these regions.

The story is very different for the integrated circuit era. Rather than broad electronics industry experience serving as an advantage, integrated circuit design and manufacturing technology and market developments make prior experience from only one type of production – planar transistors – advantageous for performance. Since relevant knowledge was located in only a few firms that kept tight control over it, access to it was available mostly to workers inside existing integrated circuit firms working at the forefront of technology. This placed existing electronics firms at a disadvantage and provided an opportunity for new firms to outperform their established counterparts. Particularly if spinoff firms are able to take advantage of technological, market and managerial knowledge gained through the founder's previous employment, spinoff firms should be more likely to out-perform both other new firms and established firms. The semiconductor market was rapidly developing both in terms of technology and economic

importance. This rapid development resulted in both strategic disagreements about technology and products and economic incentives for scientists, engineers, and managers to start their own firms and take their knowledge with them in pursuit of their own product objectives, and, perhaps, economic gains. Simply because Fairchild, one of the pioneering firms in monolithic integrated circuit development, was located in Silicon Valley and experienced events that increased incentives for employees to form spinoffs, this region became a prime candidate for substantial development during the integrated circuit era through the mechanism of spinoff firms. Through the closed and tacit nature of knowledge in the integrated circuit era, and a foundering innovative firm, Fairchild, an opportunity emerged for Silicon Valley to develop as the leading cluster for integrated circuit production.

It appears that previous experience and co-located collaboration are able to provide advantages for firm entry and performance in the transistor era while the spinoff mechanism provides advantages that allow spinoff firms to unseat established firms in the integrated circuit era. If the mechanisms that allow for firm entry and performance vary between the transistor and integrated circuit eras in this way, we should see the following:

> Corollary 1: The proportion of entrants that are spinoff firms is higher in the integrated circuit era than in the transistor era.

The chapters that follow will examine these eras using detailed firm-level data, testing the hypotheses previously presented to determine to what role each of the previously presented theories played in the development of the semiconductor industry.

Chapter 4: Research Data

This dissertation examines the development and evolution of the semiconductor industry, and thus requires industry data beginning in its formative years. This chapter discusses the identification of data sources, as well as the methodology used to collect and structure the data in a way conducive to analyses.

Firm Identification and Production Data

For any exercise in examining the development of an industry through firm entry and performance, two groups of firms must be identified: potential entrants into the industry, and the firms that decide to enter. Industry-wide publications are particularly rich sources to gather data necessary to identify such firms, and buyers' guides and other indexes have previously been used for industry studies (Buenstorf & Klepper, 2009). Additionally, the need for detailed information such as what products are produced and the location of each firm lends itself rather naturally to a buyers' guide and other industry periodicals that identify producers.

The *Electronics' Buyer's Guide (EBG)* met the data requirements and was identified as a source for firm data. The *EBG* began as a special edition of the industry periodical *Electronics*, and over time became a stand-alone annual guide for electronics buyers, with publication ceasing with the 1987 edition.[25] The *EBG* listed a large number of electronics and electronics-related

[25] While the *EBG* was an annual publication of Electronics, there does not seem to have been a 1966 publication of the guide. This determination was made by comparing the volume numbers between the 1965 and 1967 editions. As a result, no 1966 *EBG* data exists or was collected for this work.

products each year, and indicated which firms were producing each product. However, given the broad coverage of the *EBG* across a number of electronics products, proper scope had to be applied to the data collection effort to capture only semiconductor-related products. As documented in the last chapter, the transistor was the first mass-produced semiconductor device, and thus began the era of semiconductor electronics. The transistor would soon be followed by other semiconductor devices, such as the integrated circuit, co-invented by Fairchild and Texas Instruments in 1958 and first produced in 1961 (Braun & Macdonald, 1982). In examining the semiconductor industry, four semiconductor-related products are of particular interest – transistors, active modules, diodes and integrated circuits,[26] and it was for these four products that complete producer data were collected. Data on the approximately 3,000 producers of these products were gathered on an annual basis between 1949 and 1987,[27] which also allowed for the determination of how long each producer survived within the industry. Additionally, firms producing products related to transistors and integrated circuits were identified in the year preceding the introduction of these product categories in an effort to determine potential entrants.

Digitizing production and location data for almost 3,000 firms over 40 years was a fairly daunting, time-consuming task. Moreover, the effort required to build the novel dataset for this dissertation also involved collecting and structuring the data so that it could be accessed relatively easily, while at the same time maintaining backup records of the physical pages of the directory for audit and quality assurance purposes. Data collection began with scanning the

[26] Active Modules were indicated as a related product to Integrated Circuits in the *Electronics' Buyers Guide* listings of the early 1960's.
[27] 1949 is the first year that transistor producers were listed, and 1987 is the last year that the *EBG* was published.

pages from the *EBG* into an electronic format.[28] The scanned images were then processed, and, using optical character recognition (OCR), were converted to text. In the cases where scans could not be performed to an adequate quality, either because of fragile editions of the *EBG* or cases where text was located near the binding of the book, digital photographs were taken of the text that could not be scanned. *EBG* text that had been captured in this fashion was transcribed manually and merged with the OCR-converted text to form a complete record of the semiconductor producers for each year in a single text document.

With the text describing producers in separate documents for each year of production, the text had to be organized in such a way that it could be accessed for analysis. The text files were converted into a MySQL data structure using a custom program written by Jeff Plotzke, which connected each year of a firm's production into a single record and facilitated the process of consolidating entries from the same firm over time that may have had slight name variations. Following this process, the MySQL data was imported into a database consisting of three separate tables. The Products table included information about all products that were captured from the *EBG*, including the individual product categories for active modules, diodes, integrated circuits and transistors over time. For storage of firm-level data, the Operations table included the name of the firm and the location at which it operated in a given year. Given that four regions of interest had been identified using information on the previous location of electronics clusters (Boston, Los Angeles and New York City) and the prominence of Silicon Valley, the Census Bureau's definitions of Consolidated Metropolitan Areas (CMSAs) were consulted to

[28] The scans were stored as .tif image files to facilitate optical character recognition processing.

determine which firms were located within these regions. Copies of the images that were used for this purpose are included in Appendix A. Finally, the Datapoints table served to connect the Products and Operations table by denoting in which year each firm produced a given product. These three tables combined served as a complete record of the semiconductor production of approximately 3,000 semiconductor firms in the United States between 1947 and 1987.

Once the database was created, significant effort was focused on understanding the products produced by each firm. Over time, the product categories changed. Some were added, some were dropped, some changed names, and others were split into sub-categories as technology evolved. To reconcile the product classifications over time and provide a visual aid showing how the product categories evolved over time, a product "tree" of sorts was developed. The goal of the product tree was to demonstrate ties between product classifications, where possible, to simplify the many product classifications down to a set of real products or application areas over time. In order to demonstrate a connection between classifications, several methods were used. First, the names of the classifications in adjacent years were compared to determine whether the two products were similar enough to be simple name changes that occurred within the buyer's guide that were truly the same product. Next, technological knowledge was used to reconcile some of the classification changes, whether the change was a product category that was split over time or a more significant name change that was not resolved by simply comparing names. Finally, where there was doubt with either name-based evidence or technical knowledge, data examining the commonality of firms producing each of the product classifications were used in order to make a more informed decision as to whether the two product classifications were really the same product. Once the product tree was developed, there were still a large number of

product "lines" within the tree that existed over time. In an effort to provide some structure to the large number of product classifications that existed, specifically within integrated circuits, product categories were created. Using readings regarding the industry as a guide as well as engineering knowledge, eight broad product categories were created for integrated circuits: Film, Hybrid, Monolithic Custom, Monolithic Linear, Monolithic Logic, Monolithic Memories, Monolithic Microprocessor and Monolithic Other. Within each of the broad categories, there were a number of distinct product classifications that were separated into narrow product categories, largely following the category tree that was previously described, which resulted in 50 narrow product categories.

Performance Data

With data collected on all semiconductor producers, which products each produced, and where they were located over time, the next data to be collected were regarding the performance of these firms over time. Unfortunately, no unified set of performance data exist for the semiconductor industry over the time period under consideration. Therefore, data from two separate sources were used and the performance of firms in the transistor and integrated circuit eras were considered separately using these sources. For firm performance during the transistor era, data on "leading firms" in Tilton's (1971) book were used. These data appear to identify the top 10-12 transistor firms every three years between 1957 and 1966, inclusive. Due to concerns about the sales of integrated circuits during the latter half of the 1960's biasing the otherwise transistor-centric data, performance data from 1966 were omitted for the purpose of transistor firm performance analysis. The data for firm performance in the integrated circuit era were of

higher resolution, as the firm Integrated Circuit Engineering (ICE) collected annual information on the sales of U.S. semiconductor firms whose sales exceeded a non-negligible threshold between 1974 and 2002. These data, originally used in Klepper (2009a) were used to identify the top performing firms of the integrated circuit era.

Heritage Data

While the previous data collection efforts involved collecting and in some cases digitizing pre-existing information that had previously been collected by individuals or organizations, the method employed to collect information on the heritage of firms was entirely different. While some pre-existing sources of heritage information were consulted, this effort consisted of a large amount of searching through a variety of sources in order to attempt to identify firm heritage information for as many firms as possible. Ultimately, the effort resulted in a large amount of novel heritage information, especially for firms outside of Silicon Valley.

In pursuing heritage data for all firms, it seemed that there were a few natural categories of firms that provided structure to the data collection effort. Firms that were pre-existing prior to their entry into transistors and integrated circuits are considered diversifying firms, while firms that were founded in order to pursue semiconductor production were classified as new firms. Although the classification of firm heritage for all firms in the dataset was the desired result, a number of firms were not classified through these efforts, and these firms were denoted as unclassified. The classifications used are mutually exclusive and collectively exhaustive, meaning that any evidence classifying a firm as a diversifier would preclude it from being a startup, etc.

Given the structure provided within the classification system, diversifying firms seemed to be a logical place to start with the heritage data effort. First, the *EBG* was consulted to determine if the firm had previously produced products related to transistors and integrated circuits. Technical information and engineering knowledge was used to determine which products were related to transistors[29] and integrated circuits,[30] and the first year of a firm's production of each of these products was retrieved from the *EBG*. Any firm that produced a product at least five years prior to its entry in transistors or integrated circuits was listed as a producer of that product, and the product produced by the firm that was most closely related to transistors or integrated circuits was used as the classification for the firm's technological background. This portion of the analysis allowed me to account for knowledge from the production of other related products that might be beneficial to the production of transistors and integrated circuits. Additionally, the *EBG* included an index each year listing all firms that were producing electronics or electronics-related products. This index was searched for all firms in the dataset in an attempt to identify the first year that each firm produced an electronics or electronics-related product. If this year was at least 5 years prior to the firm's first transistor or integrated circuit production, the firm was classified as an electronics diversifier.[31] Finally, the last variety of diversifying firm was a firm that diversified from other non-electronic products and had produced such a product for at least

[29] Products related to transistors, listed in order of technological proximity: Vacuum Tubes, Amplifiers, Rectifiers, Switches, Resistors, Tube Parts.

[30] Products related to integrated circuits, listed in order of technological proximity: Transistors, Diodes, Active Modules.

[31] If a firm was a semiconductor diversifier, this classification took precedence over the electronics diversifier classification. Essentially, the experience that was most closely related to transistors or integrated circuits was used as the firm's only heritage.

five years prior to entry in transistors or integrated circuits. Evidence of this type of firm activity was gathered using several types of sources but was focused on identification of the date of founding or first production. Sources included general web searches, incorporation records, Dunn & Bradstreet's Million Dollar Database, Who's Who and Lexis-Nexis.

A large number of firms in the dataset were classified as diversifiers through *EBG* and other data, but there were still a number of firms that had not produced electronics or electronics-related products or been founded at least 5 years prior to their production of transistors or integrated circuits. As a result, additional background information for these firms needed to be collected, in the hopes that these firms would be able to be identified as new firms, and possibly firms that were new but were founded by individuals with previous experience in transistors or integrated circuits, which are defined as spinoffs. For transistor firms, Tilton (1971) was referenced to gather background information on a number of transistor firms. For integrated circuit firms, a total of 101 firms were featured on the ICE listings that were previously used to gather performance data, and for 92 of these firms Klepper (2009a) traced the pre-entry history of each firm, including whether it was a spinoff and if so, its "parent" firm (i.e., the prior semiconductor employer of the spinoff's founder). These two sources of data enabled the identification of the backgrounds of the largest transistor and IC producers that were not diversifiers. Additionally, a genealogy of Silicon Valley semiconductor producers compiled by the trade organization Semiconductor Equipment and Materials International (SEMI) ("Silicon Valley Genealogy Chart," 1995) was used to trace the origins of the other non-diversifiers located in Silicon Valley. These efforts exhausted all pre-existing collections of information

about new semiconductor firms and spinoffs, requiring a bit of creativity regarding further data collection.

While the *EBG* proved to be a fruitful source for information on semiconductor production, another publication, *Electronic News,* served as an excellent source for information on the founding of firms and mobility of key employees. *Electronic News* was a weekly tabloid-like publication that featured a large amount of information on rumors and developments within the industry. Every edition of *Electronic News* between 1957[32] and 1987 was examined in an attempt to identify the background of any remaining firms that had not been classified as diversifiers. Additionally, data from Who's Who, Dunn & Bradstreet's Million Dollar Database, the publication *Leaders in Electronics,* and general web searches were used to determine the founder(s) of as many non-diversifiers as possible, as well as the previous employer of each founder to determine whether the firm was a spinoff or a startup. Finally, for the remaining firms that had not yet been classified, state incorporation records were consulted. If a firm had no other supporting documentation for its heritage classification and was incorporated within 5 years of its first production of transistors or integrated circuits, the firm was classified as a startup/spinoff, essentially a firm that we knew was a new firm, but its exact origin was not known. Of the 919 transistor and integrated circuit firms, 775 were classified while 144 were not able to be classified. More detailed firm background information for transistor and integrated circuit firms is included in Chapters 5 and 6, respectively.

[32] *Electronic News* was first published in 1957.

Through a great deal of effort and creativity, the appropriate data were collected to allow for the examination of the development and evolution of the industry. While several existing data sources were collected during this effort, the combination of these sources as well as original data collected from sources such as *Electronic News*, *Electronics' Buyer's Guide* and state incorporation records resulted in the creation of a novel dataset that should prove valuable not only for this dissertation, but for significant work in the future.

Chapter 5: Transistor Entry & Performance

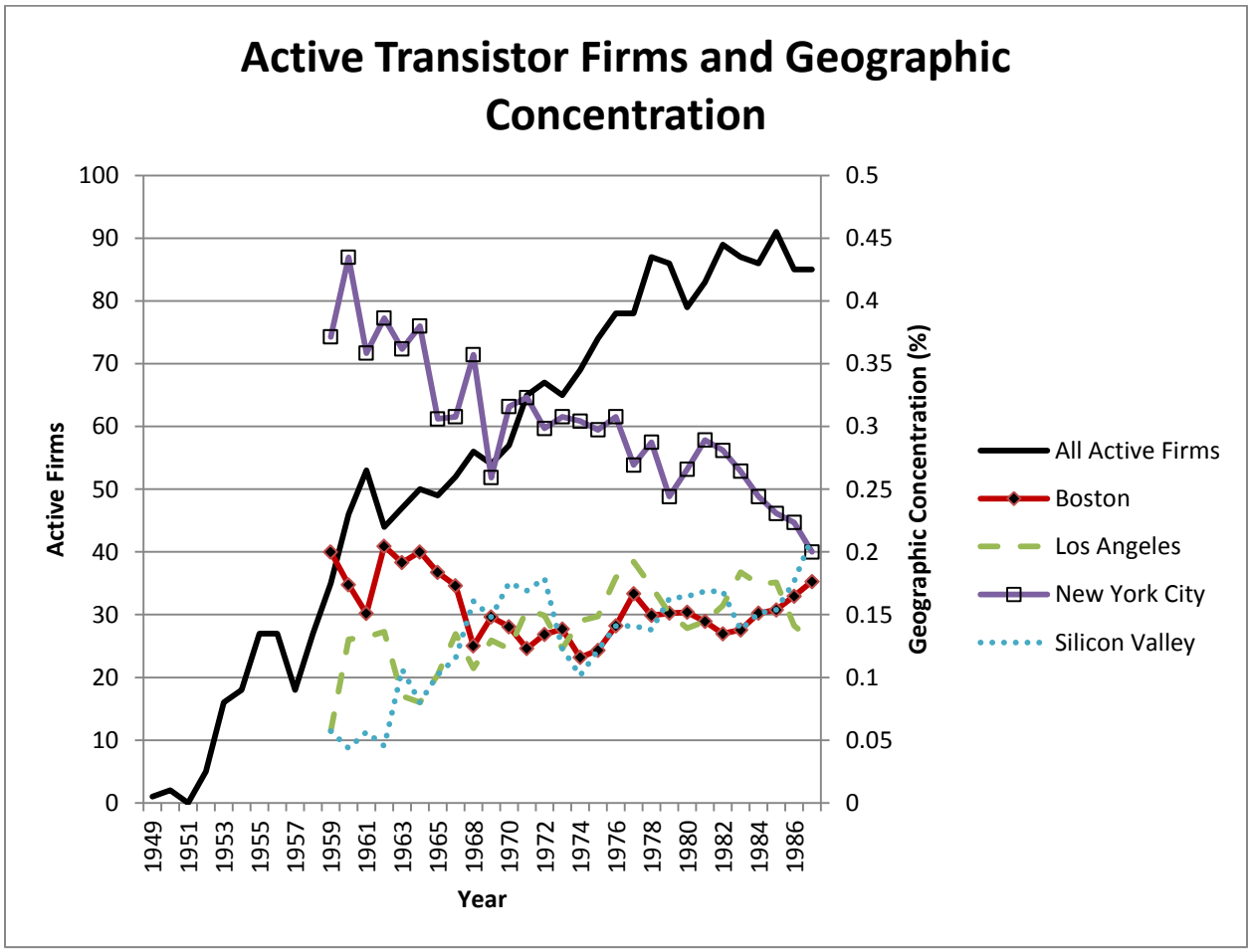
Having developed specific hypotheses reflecting the technological context of the transistor era in Chapter 3, this chapter uses the dataset described in the previous chapter to test these hypotheses. The chapter examines how firm attributes affect firm entry into transistors and firm performance following entry. The role that a firm's location and its previous production experience (in the case of diversifying firms) plays with respect to the firm's propensity to produce transistors will be examined, as well as how the same attributes, including the spinoff status of a firm, affect the firm's survival following entry. Before performing any statistical analyses, the number of entrants as well as the location and background of entrants were examined to better understand trends over time. These results will be presented in the next two sections, followed by the entry and performance analyses in the two subsequent sections.

Transistor Producer Location Distribution

As mentioned in the previous chapter, firm location data was gathered from the *EBG*, and this location was compared to the definition of the Consolidated Metropolitan Statistical Areas for Boston, Los Angeles, New York City and Silicon Valley (San Francisco) as defined by the US Census Bureau to classify firm location in terms of presence in a cluster.

With the invention of the transistor by Bell Labs in 1947, commercial production of the device commenced shortly thereafter, the first transistor producers being listed in the *EBG* in 1949. Figure 15, below, shows the total number of active transistor producers and their geographic concentrations.

Figure 15: Transistor Firm Location over Time



While the maximum number of active transistor producers (92) was reached in 1978, 53 active producers were present by 1961, showing the substantial growth in the number of producers in the first 12 years of commercial production. By 1987 there was no visible shakeout within the industry, with substantial growth in the number of producers occurring until 1977 and a fairly stable number of producers afterward. From 1959 through 1987, the New York City region had

the largest share of transistor producers,[33] with the region faring very well prior to 1959. Given the role of Bell Labs as the originator (and disseminator) of a great deal of transistor knowledge, and the organization's location in the New York City area, the prominence of the region is not surprising. The largest concentration of producers in the region occurs in 1960 (43 percent), with the concentration of firms gradually decreasing to approximately 20 percent in 1987. The other regions had a much smaller share of producers in 1959, with Boston having 19 percent and Los Angeles and Silicon Valley having 6 percent. Over time, the concentration of producers in these regions fluctuated a bit, with each of the three regions converging to a concentration of approximately 20 percent by 1987. Overall, New York City loses a great deal of producer concentration, with each of the other regions (including outside of the clusters) achieving approximately 20 percent share of producers by 1987.

Transistor Producer Background Distribution

Following the examination of the number and location of transistor producers over time, the technological backgrounds were examined next. To determine these backgrounds, several sources of data were used, as described in the previous chapter. The results of this substantial data effort are shown below in Table 4 as firm background classifications for all transistor firms. As can be observed, a large portion transistor producers was classified through this effort.

[33] The number of transistor producers is less than 30 prior to 1959, meaning that the geographic concentrations vary wildly as the mix of producers changed each year. As a result, geographic concentrations are reported starting in 1959.

Table 4: Transistor Firms by Background and Region, 1949-1987

Location	Amplifiers	Diversifiers	Electronics	New	Other[34]	Tubes	Unknown	Total	Spinoffs
Boston	3	7	0	8	4	3	4	29	4
Los Angeles	3	17	5	13	3	0	11	52	5
New York City	2	17	17	22	7	7	34	106	10
Silicon Valley	0	2	1	26	2	0	3	34	18
Other	4	20	22	23	9	3	17	98	11
Total	12	63	45	92	25	13	69	319	48

Transistor producers are primarily composed of new firms (29 percent), diversifiers (20 percent), and electronics producers (14 percent), with amplifier, tube and other electronics producers accounting for less than ten percent each of transistor entrants. The remaining (22 percent) are unclassified firms. Overall, the distribution of technological backgrounds is fairly similar between regions, with outliers including new firms and Silicon Valley (76 percent of Silicon Valley entrants are new firms compared to 29 percent of overall entrants), and unknown firms in New York City (32 percent of New York City entrants compared to 22 percent of overall entrants). Additionally, the timing of entrants is distributed fairly evenly over the decades, with 80 entering by 1960, 95 entering between 1961 and 1970, 93 entering between 1971 and 1980 and 51 entering between 1981 and 1987.

Firm Entry Analysis

The first issue considered is the rate at which pre-existing electronics firms diversified into transistors, examining specifically which firm attributes may have influenced both the decision

[34] The Other category includes firms producing active modules, capacitors, diodes, filters, rectifiers, resistors, switches and tube parts, as each of these categories did not have many entrants.

to diversify into transistors and the speed at which this diversification occurred. This analysis is motivated by theory that suggests that firms with previous experience are most likely to begin producing related products (Scott-Morton, 1999; Klepper & Simons, 2000; Lane, 1989), contextualized to the transistor era in the following hypothesis generated in Chapter 3:

Hypothesis 1: Among existing firms at the time that transistors are invented, firms with related electronics experience, specifically tube manufacturers (who faced possible obsolescence from the introduction of the transistor), are more likely to begin producing transistors, all else held constant.

To test this hypothesis, a series of Cox Proportional Hazard Regressions were performed to examine the effects of specific variables on the amount of time until a firm began producing transistors. Controls for location were included to account for any region-level effects that might be present within either the existing electronics clusters (Boston, Los Angeles or New York City) or the semiconductor cluster that would emerge in future years (Silicon Valley). After examining the product categories that were present in the *EBG* in 1948, several related product categories were identified due to similar functions of the product with respect to transistors. Specifically, amplifiers, rectifiers, resistors, switches, tubes and tube parts were considered related product categories. The item produced by the firm that was most closely related to transistors was used as the basis for classifying the technological background of the firm. As described in chapter 3, transistors were seen as a direct replacement for vacuum tubes, and thus tubes were considered first for the classification of previous firm experience. Following tubes, devices that performed functions that transistors were capable of performing were considered next, including amplifiers, rectifiers and switches. Finally, other products that were related to transistor applications and tube production, such as resistors and tube parts, were considered.

Given that all firms were considered starting in 1948, no time-based variables were included in the analysis.

Annual observations were created for each firm that was producing related products in 1948. Once the firm began producing transistors, the firm was marked as having "failed" for the purposes of the proportional hazard survival regression, and observations were not generated for future years. For firms that never began producing transistors, observations continued through 1987 as that was the final year that production data existed for transistors and these firms were treated as censored.

The first regression includes all firms in related industries in 1948, the year prior to the first listing of transistors in the *EBG*.[35] Potential entrants included 197 amplifier firms, 24 rectifier firms, 49 resistor firms, 135 switch firms, 89 tube firms, and 45 tube part firms, for a total sample of 539 potential entrants. Of the 539 potential entrants, 21 eventually produced transistors. The results of the regression are shown below in Table 5 and are reported as hazard coefficients, meaning that positive (negative) numbers have a positive (negative) effect on firm entry.

[35] The three firms that entered transistors less than four years prior to entering integrated circuits are omitted from this analysis as they are not treated as transistor firms in our background classifications.

Table 5: Transistor Entry Analysis

	N	M1	M2
Location			
Boston	41	0.50	0.34
		(0.80)	(0.80)
Los Angeles	34	-0.03	-0.09
		(1.06)	(1.07)
New York City	223	0.40	0.29
		(0.49)	(0.49)
Silicon Valley	10	1.19	0.61
		(1.06)	(1.07)
Background			
Amplifiers	197		0.88
			(1.07)
Switches	135		0.04
			(1.22)
Tubes	89		2.07**
			(1.05)
Resistors	49		0.35
			(1.41)
Log Likelihood		-130.93	-124.61
Observations	539	539	539

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

First, the role of geography alone was examined to determine if existing firms located where the semiconductor industry was clustered, specifically Boston, New York and Los Angeles, or where it was going to be clustered, in Silicon Valley, were more likely to begin producing transistors. Model 1 includes separate dummies for each region, with the remaining United States serving as the omitted category. As can be observed, no coefficient estimates are statistically significant, indicating that firms located throughout the United States were just as likely to begin producing transistors, regardless of their location.

To directly test Hypothesis 1, the role of technological background was examined to determine if firms in related industries were more likely to begin producing transistors. Building on Model 1, Model 2 also includes dummies for geography and technological background. The omitted categories for location and background variables are firms located outside of the industry clusters and tube part firms, respectively. When controlling for both geography and technological background, the only statistically significant effect is for tube firms, with a p-value of 0.014 and a positive coefficient estimate, indicating that firms with previous vacuum tube experience were more likely to enter transistors, compared to firms with other experience, regardless of location. Given the statistical significance of the tube background coefficient estimate, we can reject the null hypothesis that firms with related electronics experience, specifically tube manufacturers, are no more likely to begin producing transistors than other firms, all else held constant.

To summarize, the results of transistor entry regressions indicate that when accounting for both geography and technological background, it was technological background alone that influenced the decision to enter transistors. Specifically, firms that were producing vacuum tubes were more likely to begin producing transistors than firms with other backgrounds, regardless of location. As discussed previously in chapter 3, this may have been either because of the fear that transistors would cannibalize the vacuum tube market rather quickly, or the realization by these tube producers that they possessed many of the key complementary assets which could be used in the production and sales of transistors. Motivation based on previous production seems likely to encourage vacuum tube firms, many of which had been located in existing electronics clusters, to transition to producing transistors, creating a relationship between the location of vacuum tube firms and transistor firms, at least as far as transistor entrants are concerned.

Firm Performance Analysis

While we are interested in what influences firms to enter a given industry, we are also interested in what leads firms to perform well following entry. As discussed previously, there are three separate mechanisms within the transistor era that may provide advantages to firms following their entry into transistor production: previous production in related products, first-mover advantage, and knowledge spillovers through agglomeration economies. Following previous practice in the literature, firm performance will be analyzed, using two measurements: survival and market share attainment (Klepper, 2007a, 2009a).

Given that a set of hypotheses related to performance will be examined using both survival and market share attainment data, it seems appropriate to re-introduce the hypotheses at a broad level before testing them. As we saw in the entry analysis for transistors, previous experience, particularly as a vacuum tube producer, was influential in the decision to begin producing transistors. Likewise, we expect that this previous experience may have an effect on firm performance following entry, as has been shown in a number of studies (Klepper & Simons, 2000; Lane, 1989). As a result, we can test the following hypothesis regarding the influence of previous experience by examining whether the coefficient estimates for previous backgrounds are statistically significant.

Hypothesis 2: Among transistor producers, firms with related electronics experience, specifically tube manufacturers, will perform better than other firms, all else held constant.

Additionally, there is some thought that early producers of transistors would be able to enjoy first-mover advantages allowing such firms to gain knowledge and experience regarding

transistor manufacturing. This could provide the firm with advantages regarding profitability or market share, placing the firm at an advantage with respect to other firms that began producing transistors later. The possible existence and influence of a first-mover advantage can be examined and the hypothesis below tested by controlling for the entry date of the firm to transistor production and examining whether the coefficient estimates for this variable is statistically significant.

Hypothesis 3: Among transistor producers, firms that begin producing transistors earlier will perform better than other firms, all else held constant.

Yet another theory indicates that co-located firms may enjoy advantages with respect to knowledge spillovers. This should give these firms an advantage with respect to firms located outside of such clusters. While there is some doubt that these effects would be notable during this era given the openness of and efforts to disseminate knowledge by Bell Labs all around the US, we should be able to tease out whether such an effect did exist and how influential it might have been during the transistor era. Using the industry data collected, we can examine the possible existence and influence of agglomeration economy effects and test the hypothesis below by controlling for the location of the transistor producers and examining whether the coefficient estimates for each cluster are statistically significant.

Hypothesis 4: Among transistor producers, firms located in Boston, Los Angeles and New York will perform better than firms located elsewhere, all else held constant.

If firms experience advantages by being located in a cluster, one might also expect that some firms previously operating elsewhere would decide to establish their novel transistor operations

in a cluster. We should be able to see the effects of such a move within the data, allowing us to test the hypothesis below.

Hypothesis 5: Among transistor producers with prior electronics production, firms that locate their transistor production in a different region that previous production and in the Boston, Los Angeles, and New York regions are more likely to out-perform other comparable firms, all else held constant.

The previous hypotheses will be examined using both survival and market share attainment data.

Survival Analysis

For this analysis, survival is considered only for transistor production. Therefore, no other production was considered, even if the firm was active in other semiconductor areas. A Cox Proportional Hazard Model was used to examine how several attributes contributed to the survival of transistor producers. Dummy variables representing the various firm backgrounds are included in the analysis to examine the role of firm backgrounds on survival. To examine the possible first-mover benefits for transistor survival, a variable indicating the year of first transistor production was included. To account for any region-level effects that might be present within either the existing electronics clusters (Boston, Los Angeles or New York City) or the semiconductor cluster that would emerge in future years (Silicon Valley), dummy variables were included indicating the location of the firm.

All firms that began producing transistors prior to 1964[36] were included in the analysis, which included 118 transistor producers in total. These firms included 49 new and unclassified firms, 20 diversifiers, 24 electronics firms, 8 firms producing other electronics,[37] 6 amplifier firms and 9 tube firms. In terms of firm location, a large number of transistor producers (47) were located in the New York City region, with 13 firms located in Los Angeles, 13 in Boston and 7 in Silicon Valley, and the remaining firms located outside of the industry clusters. The results of the Cox Proportional Hazard Regressions are shown below in the Table 6 and are reported as hazard coefficients. While we are examining the survival of transistor firms, the analysis measures the hazard of a firm exiting production of transistor products, meaning that positive (negative) numbers have a negative (positive) effect on firm survival since they have a positive (negative) effect on firm exit.

[36] The use of 1964 as a cut-off date is driven by two factors. First, the integrated circuit is first listed in the *EBG* in 1965, and it is desirable to make sure that integrated circuit production is not influencing the survival of transistor firms. Additionally, the final market share attainment datapoint that will be used for the next analysis is in 1963. In the interest of consistency, the cut-off of prior to 1964 is used to account for the impending introduction of the integrated circuit and keep the sample consistent between both the survival and market share attainment analyses.

[37] Due to the small number of firms producing each product category, capacitors, diodes, filters, rectifiers and switches are considered "other electronics".

Table 6: Transistor Survival Analysis

	N	M1	M2	M3
Location				
Boston	13	-0.21	-0.19	-0.22
		(0.35)	(0.36)	(0.36)
Los Angeles	13	-0.16	-0.24	-0.27
		(0.34)	(0.35)	(0.36)
New York City	47	0.05	0.14	0.12
		(0.23)	(0.24)	(0.24)
Silicon Valley	7	-0.92*	-0.91*	-0.93*
		(0.49)	(0.50)	(0.50)
Relocating Firm	6			-0.39
				(0.49)
Background				
Amplifiers	6		-0.29	-0.21
			(0.49)	(0.50)
Diversifiers	18		0.05	0.05
			(0.30)	(0.30)
Electronics	24		0.16	0.22
			(0.26)	(0.27)
Tubes	9		-0.82*	-0.82*
			(0.46)	(0.46)
Other Electronics	8		-0.51	-0.45
			(0.44)	(0.45)
Entry Year	118	0.07**	0.04	0.04
		(0.03)	(0.03)	(0.03)
Log Likelihood		-426.53	-423.48	-423.12
Observations	118	118	118	118

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

The analyses begins by looking only at the role of geography was examined to determine if transistor firms were more likely to survive where the semiconductor industry (and previous electronics firms) was clustered, specifically Boston, New York, Los Angeles, and Silicon Valley. Firms located outside any of the clusters served as the omitted category. The initial results, presented in column M1, show that the Silicon Valley region had a coefficient estimate that was negative and statistically significant. This indicates that, controlling for time of entry, firms located in Silicon Valley were more likely to survive compared to firms located elsewhere.

The coefficient estimate for year of entry is statistically significant and positive, indicating that firms that entered later were more likely to fail. The inverse of this statement is that firms that entered earlier were less likely to fail, indicating an advantage for early entrants, all else held constant.

Next, the role of technological background was examined jointly with location to determine if firms with experience in related industries were more likely to survive following their entry into transistors. This is marked as Model 2 in Table 6. The Silicon Valley regional dummy maintains its significance and relative magnitude, while only the vacuum tube background dummy is statistically significant and negative. The significance of the entry year variable was eliminated, suggesting there was no first mover advantage with respect to survival once firm background was considered. These results indicate that firms located in Silicon Valley and firms that previously produced vacuum tubes prior to transistor entry were more likely to survive compared with other similar firms.

The results of these regressions suggest that firms located in Silicon Valley appeared to have a regional advantage. However, this apparent regional effect is driven heavily by the presence of one distinct firm – Fairchild – which is influential in the analysis of the Silicon Valley effect because of the rather small number of firms located there (7). Additionally, the significance of the first-mover advantage coefficient estimate was eliminated once firm background was included in the model. Technological background, however, specifically a background in vacuum tube production appears to have been advantageous with respect to survival. Thus, we find consistent support for only Hypothesis 2 (background advantage). We are unable to find support for Hypothesis 4 regarding regional effects in Boston, Los Angeles and New York, as

none of the coefficient estimates for these regional variables are statistically significant. In summary, it appears that heritage and location played an important role in how well transistor producers were able to survive, while the first-mover advantage was not overly important.

Market Share Attainment

While survival is one measure of firm performance, it is by no means an ideal measure. While survival will inform as to how long a firm was able to continue producing products, it does not convey the performance of the firm with respect to whether it was able to lead the industry in sales. However, it is not always easy to collect market share or sales data, causing a number of analyses to rely on survival data. To allay the concerns with survival data, data regarding firm market share attainment were gathered, and used to complement the previous survival analysis.

The top 10-12 market leaders for transistors were identified in Tilton's (1971) book in three year increments between 1957 and 1963.[38] A logit model was used to examine how several attributes contributed to the market share attainment of transistor producers, with one observation for each firm included in the analysis. The dependent variable was the time required for the firm to achieve market leader status according to Tilton (1971), with firms that did not achieve leader status by 1963 treated as censored. The same set of variables from the survival analysis is included in this market share attainment analysis, allowing us to test the same hypotheses with respect to market share attainment.

[38] As mentioned previously, the 1966 data, while available, were not used as a result of concerns regarding integrated circuit sales data influencing these market share rankings.

The sample of firms was the same as the previous analysis, and therefore descriptive statistics regarding the sample will not be repeated here. The results of the logit regressions are shown below in the Table 7 and are reported as coefficients, meaning that positive (negative) numbers have a positive (negative) effect on firm market share attainment.

Table 7: Transistor Market Share Attainment Analysis

	N	M1	M2	M3
Location				
Boston	13	-0.41	-0.51	-0.30
		(1.01)	(1.27)	(1.30)
Los Angeles	13	0.80	1.26	1.49
		(1.06)	(1.17)	(1.22)
New York City	47	-1.58	-3.24*	-3.27
		(1.20)	(1.89)	(2.03)
Silicon Valley	7	2.48	3.12*	3.45*
		(1.61)	(1.76)	(1.86)
Relocating Firm	6			1.67
				(1.72)
Background				
Amplifiers	6		1.14	0.92
			(1.45)	(1.43)
Diversifiers	18		-0.28	-0.27
			(1.13)	(1.14)
Electronics	24		0.08	-0.08
			(1.44)	(1.47)
Tubes	9		3.22*	3.41*
			(1.81)	(1.94)
Other Electronics	8		0.41	-0.26
			(1.47)	(1.73)
Entry Year	118	-0.62***	-0.65***	-0.69***
		(0.20)	(0.22)	(0.24)
Constant		1,218***	1,271***	1,358***
		(390.7)	(441.2)	(477.9)
Log Likelihood		-24.27	-21.65	-21.20
Observations		118	118	118

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

First, the role of geography was examined to determine if transistor firms were more likely to attain leading status where the semiconductor industry (and previous electronics firms) was clustered, specifically Boston, New York, Los Angeles, and Silicon Valley. Separate dummies were included for each region, with the remaining United States serving as the omitted category.

None of the regional variables had a statistically significant coefficient estimate, indicating that firms located in clusters were no more likely to attain leading status than firms located elsewhere. The coefficient estimate for year of entry is statistically significant and negative, indicating that firms that entered earlier were more likely to attain leading status. This result provides support for a first-mover advantage, but could also be explained by a lower level of competition early in the transistor era as the product was developing which allowed firms to more easily attain market leadership status.

Next, the role of technological background was examined jointly with location to determine if firms with experience in related industries were more likely to attain market leadership status following their entry into transistors, marked Model 2. The New York and Silicon Valley coefficient variables are negative and positive, respectively and are both significant at the 10 percent level, indicating that firms in Silicon Valley were more likely to attain leading market share status, and less likely in New York City. This result, however, does not provide support for Hypothesis 4 as the benefits were predicted to occur in Boston, Los Angeles and New York City. Additionally, the tube coefficient estimate is positive and statistically significant, which indicates that firms with previous vacuum tube experience were more likely to attain leading market share status. This finding provides support for Hypothesis 2. Additionally, the significance and approximate magnitude of the entry year variable is maintained, indicating further support for a first-mover advantage.

Finally, the dummy representing existing firms that located transistor production within an industry cluster rather than near the firm's previous production outside of a cluster was included to test Hypothesis 5. The significance of the New York coefficient estimate is eliminated, but

otherwise all other significance levels are maintained as compared to Model 2. Additionally, the coefficient estimate of the relocation variable added in Model 3 is not statistically significant, which does not provide support for Hypothesis 5

The results of both analyses indicate some mixed support for some of the hypotheses. Specifically the survival analysis shows consistent support for Hypothesis 2 (background advantage), while the market share attainment analysis demonstrates support for Hypothesis 2 and Hypothesis 3 (first-mover advantage). Hypothesis 4, related to regional advantages in Boston, Los Angeles and New York City, was not supported in either analysis. The apparent Silicon Valley effect that was evident in both analyses is driven by the existence of a single firm – Fairchild – in the region. In fact, if Fairchild is excluded from the market share attainment analysis, the regional coefficient cannot be estimated as no other firms achieve leading market share status.

In summary, it appears that background was an important attribute with respect to survival, but that a first-mover advantage explains a large proportion of the firms that achieve leadership status in the transistor era. Given that most of the first-movers were vacuum tube firms that faced possible obsolescence if they did not produce transistors, the analysis suggests that the location of the semiconductor industry during the transistor era was closely related to its location during the vacuum tube era. Thus, the transistor era only served to reinforce the location of the electronics industry in regions such as Boston, Los Angeles and New York City.

A summary of the hypotheses tested and results found in this chapter are shown in the table below.

Table 8: Summary of Transistor Results

Hypothesis	Results (+/-/No support)
Hypothesis 1: Among existing firms at the time that transistors are invented, firms with related electronics experience, specifically tube manufacturers who faced possible obsolescence from the introduction of the transistor, are more likely to begin producing transistors, all else held constant.	+
Hypothesis 2: Among transistor producers, firms with related electronics experience, specifically tube manufacturers, may be more likely to perform better than other firms, all else held constant.	+
Hypothesis 3: Among transistor producers, firms that begin producing transistors earlier are more likely to perform better than other firms, all else held constant.	+
Hypothesis 4: Among transistor producers, firms located in Boston, Los Angeles and New York are more likely to perform better than firms located elsewhere, all else held constant.	No Support
Hypothesis 5: Among transistor producers with prior electronics production, firms that locate their transistor production in a different region that previous production and in the Boston, Los Angeles, and New York regions are more likely to out-perform other comparable firms, all else held constant.	No Support

Chapter 6: Integrated Circuit Entry & Performance

Building on the analysis performed in Chapter 5, this chapter examines and tests the hypotheses that were developed in Chapter 4 for the integrated circuit era. Broadly, this chapter examines how firm attributes affect firm entry into integrated circuits and firm performance following entry. As noted in Chapter 4, important differences in results between the two eras are expected to be found. The role that a firm's location and its previous production experience (in the case of diversifying firms) in whether the firm entered integrated circuits were examined, as was how similar attributes, including the spinoff status of a firm, affect the firm's survival following entry. Before performing any statistical analyses, the number of entrants as well as the location and background of entrants were examined to better understand trends over time. These results will be presented in the next two sections, followed by the entry and performance analyses in the two subsequent sections.

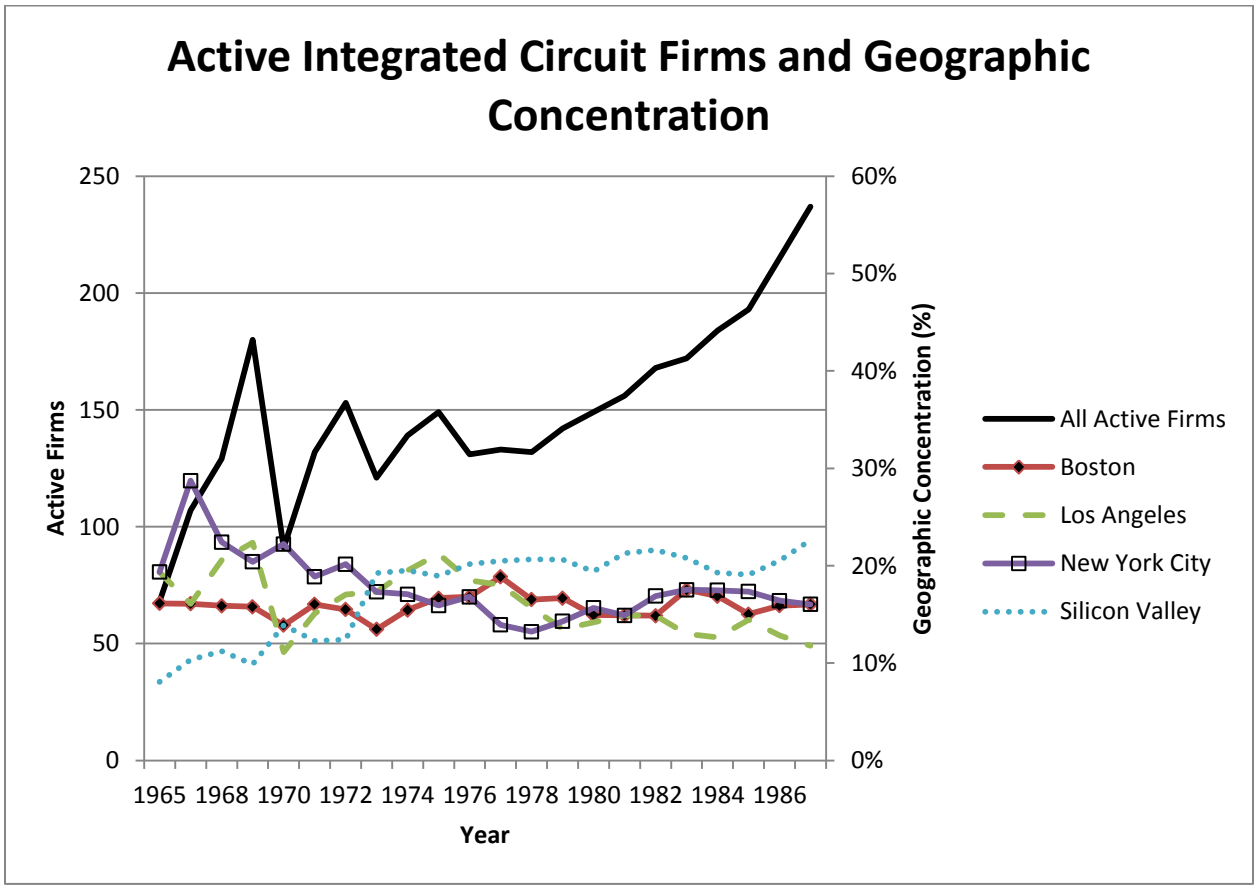
Integrated Circuit Producer Location Distribution

As mentioned previously, firm location data were gathered from the *EBG*, and the firm's location was compared to the definition of the Consolidated Metropolitan Statistical Areas for Boston, Los Angeles, New York City and Silicon Valley (San Francisco) as defined by the US Census Bureau in order to classify firm location.

Integrated circuits see a significant growth in the number of producers within the first decade, much like transistors. Following the co-invention of the monolithic integrated circuit in 1959 by Fairchild and Texas Instruments, the commercial production of the device began in 1961

(Lécuyer, 2006), with the first production recorded in the *EBG* in 1965.[39] Figure 16, below, shows the total number of active integrated circuit producers and geographic concentration of producers.

Figure 16: IC Firm Location over Time



[39] The *EBG* lists a category entitled "Active Modules" which is lists as the predecessor to the integrated circuit. However, the number of Active Module firms that go on to produce integrated circuits is very low, leading me to exclude Active Modules from the integrated circuits discussion and analysis.

The number of active integrated circuit producers more than doubled from its initial level of 63 to 133 just four years later. While the number of active producers fluctuated a great deal in the first decade, the number of producers grew on a fairly consistent basis from 1975 onward, with the maximum number of active producers occurring in 1987 (215).[40] Unlike transistors, no region consistently had the largest concentration of producers for integrated circuits. Rather, Boston, Los Angeles and New York City had fairly similar producer concentrations (between 20 and 30 percent) between 1965 and 1987, with these concentrations remaining fairly constant. Silicon Valley had a very low producer concentration in 1965 – approximately 10 percent. However, this concentration grew over time and Silicon Valley overtook the other regions in terms of producer concentration in 1976. Silicon Valley remained the leading region from 1976 onward, with a producer concentration in the upper 20 percent range in the late 1970s and early 1980s, reaching 30 percent in 1987. Overall, the main story of integrated circuits is the emergence of Silicon Valley as a leading region of producers beginning in the 1970s.

Integrated Circuit Producer Background Distribution

Following the examination of the number and location of integrated circuit producers over time, the technological backgrounds of integrated circuit producers was examined next, relying on several sources of data, as described previously. The results of this substantial data effort are shown below in Table 8 as firm background classifications for all integrated circuit firms. Overall, a large portion of integrated circuit producers was able to be classified.

[40] It should be noted that the data end in 1987 – there is no indication whether the number of producers continues growing after this point.

Table 9: Integrated Circuit Firms by Background and Region, 1965-1987.

Location	Active Modules	Diodes	Diversifiers	Electronics	Transistors	New	Unclassified	Total	Spinoff
Boston	9	6	14	13	5	28	2	77	8
Los Angeles	6	2	17	22	8	24	17	96	9
New York City	9	6	20	25	11	22	21	114	8
Silicon Valley	1	2	3	6	5	73	0	90	54
Other	18	7	33	62	15	53	35	223	24
Total	43	23	87	128	44	200	75	600	103

Integrated circuit producers are primarily composed of new firms (33 percent), electronics (20 percent), diversifiers (14 percent), and unknown firms (13 percent), with active module firms (9 percent), transistor firms (7 percent), and diode firms (4 percent) making up the remaining technological backgrounds. The distribution of technological backgrounds between regions is fairly similar with one exception – Silicon Valley. 80 percent of Silicon Valley entrants are new firms as compared to the 33 percent of overall entrants that are new firms. As a result, the shares of Silicon Valley firms with other background classifications are significantly lower than the shares in other regions. Through these data efforts, the vast majority of these new firms in Silicon Valley was found to be spinoff firms – a fact that will be important later in the data analysis when the survival and performance differential for high quality spinoff firms is examined. In contrast to transistors, the timing of integrated circuits entrants is fairly front-loaded with more than half of entrants (56 percent) entering within the first eight years (by 1973).

Firm Entry Analysis

First, the rate at which pre-existing firms diversified into integrated circuits was considered, testing the following hypothesis:

> **Hypothesis 7: Among existing firms at the time that integrated circuits are invented, firms with related electronics experience are more likely to begin producing integrated circuits than firms without electronics experience, all else held constant.**

To examine this, a series of Cox Proportional Hazard Regressions were performed to look at the effects of specific background variables on the amount of time until a firm began producing integrated circuits. Controls for location were included to account for any region-level effects that might be present within the existing transistor clusters (Boston, Los Angeles, New York City or Silicon Valley).

By examining the product categories present in the *EBG* in 1964, several product categories were identified as being similar in terms of product functionality with respect to integrated circuits. Specifically, active modules, diodes and transistors were considered related product categories. The item produced by the firm that was most closely related to integrated circuits was used as the classification for the firm's technological background. Given the similarities between many types of integrated circuits and transistors, this product category was considered first, followed by diodes which were typically constructed using semiconductor materials, followed by active modules, which often combined conventional components into a single device. Given that all firms were considered starting in 1964, no time-based variables were included in the analysis.

Annual observations were created for each firm that was producing related products in 1964. Once the firm began producing integrated circuits, the firm was marked as having "failed" for the purposes of the proportional hazard survival regression, and observations were not generated for future years. For firms that never began producing integrated circuits, observations continued through 1987 as that was the final year that production data existed for integrated circuits and these firms were treated as censored.

The sample for this analysis included all firms in related industries in 1964, the year prior to the first listing of integrated circuits in the *EBG*. Potential entrants included 67 diode firms, 51 transistor firms, and 213 active module firms, for a total sample of 331 potential entrants. Of the 331 potential entrants, 37 eventually produced integrated circuits. Transistor firms were split into germanium and silicon transistor firms, since most integrated circuits between 1965 and 1987 were produced using silicon. While the hypotheses specifically call out planar transistor production, EBG product data did not specify whether production was planar. Given that only silicon can be used for planar-based production and the popularity of planar-based transistors, silicon transistor production was used as a proxy for planar transistor production. As a result, the expectation is that silicon transistor firms may have an advantage over germanium transistor firms since they were familiar with the manufacturing processes that would later be used for integrated circuits. The results of the regression are shown in the Table 9, below, and are reported as hazard coefficients, meaning that positive (negative) numbers have a positive (negative) effect on firm entry.

Table 10: Integrated Circuit Entry Analysis

	N	M1	M2
Location			
Boston	44	0.16	-0.06
		(0.32)	(0.33)
Los Angeles	56	-0.03	0.04
		(0.34)	(0.34)
New York City	90	-0.32	-0.39
		(0.31)	(0.30)
Silicon Valley	15	0.75	0.72
		(0.48)	(0.49)
Background			
Diodes	67		-0.34
			(0.30)
Germanium Transistors	2		0.62
			(0.73)
Silicon Transistors	45		1.27***
			(0.27)
Log Likelihood		-375.38	-361.86
Observations	331	331	331

Standard Errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

First, the role of geography was examined to determine if firms were more likely to enter where the semiconductor industry was clustered at the time integrated circuits were introduced, specifically Boston, New York and Los Angeles, or where it was going to be clustered, in Silicon Valley. Separate dummies were included for each region, with the remaining United States serving as the omitted category. None of the dummies have coefficient estimates that are statistically significant, indicating that firms located in transistor clusters were no more likely to enter integrated circuits than firms in other regions.

Geography and technological background were examined jointly in Model 2, which includes a set of dummy variables for location as well as technological background. Transistor experience

was split between those firms that produced any silicon transistors[41] and firms that only produced germanium transistors, with diodes being the other included technological background variable. The omitted categories for location and background variables are firms located outside of the industry clusters and active module firms, respectively. The results in terms of the location dummies are very similar to Model 1, with none of the coefficient estimates being statistically significant. For the technological background variables, only the Silicon Transistor variable has a statistically significant coefficient estimate, which is positive, indicating that firms with previous silicon transistor experience were more likely to enter integrated circuits than similar firms with other experience.

To summarize, the results of integrated circuit entry regressions indicate that when accounting for both geography and technological background it was technological background alone that influenced the decision to enter integrated circuits. This is a similar result to the transistor entry analysis, and Hypotheses 1 (background advantage) and 6 (background advantage) are supported by these analyses.

Firm Performance Analysis

While we are interested in what influences firms' decisions to enter a given industry, we are also interested in what leads firms to perform well following entry. As discussed previously, there are three mechanisms that may provide advantages to firms following their entry into transistor

[41] Silicon transistors were more likely to have utilized some of the same manufacturing processes as would be used for monolithic integrated circuits than germanium transistors, which is why the two types of products were separated.

production: previous production of related products, knowledge gained through the spinoff mechanism, and knowledge spillovers through agglomeration economies. The possible existence and influence of two of these advantages within the context of the IC era will be examined in this section, using two measurements of firm performance: survival and market share attainment. Whether knowledge is transferred from parent to spinoff and potential effects of this mechanism will be examined in depth in the next chapter, and so this aspect will not be included in the following analyses.

As we saw in the entry analysis for integrated circuits, previous experience as a silicon transistor producer was influential in the decision to begin producing integrated circuits. Likewise, we expect this experience to have an effect on firm performance following entry, as has been shown in a number of studies (Klepper & Simons, 2000; Lane, 1989). As a result, we can test the following hypothesis regarding the influence of previous experience by examining whether the coefficient estimates for previous backgrounds are statistically significant.

Hypothesis 12: Among integrated circuit entrants with previous experience, firms that had previous planar transistor production experience were more likely to perform better than other integrated circuit producers, all else held constant.

As was the case when we examined transistors, agglomeration theory indicates that co-located firms may enjoy advantages with respect to firms located outside of such clusters. We can examine the possible existence and influence of agglomeration economy effects and test the hypothesis below by controlling for the location of the integrated circuit producers and examining whether the coefficient estimates for each cluster is statistically significant.

Hypothesis 8: Among integrated circuit producers, firms located in Boston, Los Angeles, New York and Silicon Valley are more likely to out-perform comparable firms located elsewhere, all else held constant.

If firms experience advantages by being located in a cluster, one might also expect that some firms previously operating elsewhere would decide to establish their novel integrated circuit operations in a cluster. We should be able to see the effects of such a move within the data, allowing us to test the hypothesis below.

Hypothesis 9: Among integrated circuit producers with prior electronics production, firms that locate their integrated circuit production in a different region that previous production and in the Boston, Los Angeles, New York and Silicon Valley regions are more likely to out-perform other comparable firms, all else held constant.

Using production and performance data in a similar fashion to the examination that was performed for transistor firms, these three hypotheses will be examined below.

Survival Analysis

Survival is considered as the time that the firm was able to produce integrated circuits, the same procedure adopted when looking at survival in transistor production. A Cox Proportional Hazard Model was used to examine how several attributes contributed to the survival of integrated circuit producers. Dummy variables representing the various firm backgrounds are included in the analysis to examine the role of firm backgrounds on survival. To account for any region-level effects that might be present within either the existing transistor clusters (Boston, Los Angeles, New York City or Silicon Valley), dummy variables were included indicating the

location of the firm. Finally, to control for possible advantages firms many have enjoyed from moving to a cluster, a dummy was included to account for this firm activity.

All integrated circuit firms that began producing integrated circuits prior to 1987[42] were included in the analysis, which totaled 575 integrated circuit producers. These firms included 263 new or unclassified firms, 122 electronics firms, 84 diversifiers, 42 active module firms, 43 transistor firms, and 21 diode firms. In terms of the location distribution of the firms, a large number of integrated circuit producers (111) were located in the New York City region, with 94 firms located in Los Angeles, 80 in Silicon Valley and 75 in Boston, with the remaining firms located outside of the industry clusters. The results of the Cox Proportional Hazard Regressions are shown below in Table 10 and are reported as hazard coefficients. While we are examining the survival of integrated circuit firms, the analysis measures the hazard of a firm exiting production of integrated circuit products, meaning that positive (negative) numbers have a negative (positive) effect on firm survival since they have a positive (negative) effect on firm exit. To control for possible first-mover advantages, the entry year for each firm is also included in the regression.

[42] Firms that entered in 1987 were not able to exit during the time when data were collected and thus these 25 firms were omitted from this survival analysis.

Table 11: Integrated Circuit Survival Analysis

	N	M1	M2	M3
Location				
Boston	75	-0.26	-0.23	-0.23
		(0.17)	(0.17)	(0.17)
Los Angeles	94	0.09	0.05	0.06
		(0.14)	(0.14)	(0.14)
New York City	111	0.09	0.12	0.12
		(0.13)	(0.13)	(0.13)
Silicon Valley	80	-0.34**	-0.32*	-0.32*
		(0.17)	(0.17)	(0.18)
Relocating Firm	13			0.02
				(0.32)
Background				
Active Modules	42		0.10	0.10
			(0.19)	(0.19)
Diodes	21		-0.48	-0.48
			(0.30)	(0.30)
Electronics	122		0.03	0.03
			(0.13)	(0.13)
Transistors	43		-0.79***	-0.79***
			(0.22)	(0.22)
Diversifiers	84		0.11	0.11
			(0.15)	(0.15)
Entry Year	575	-0.03***	-0.03***	-0.03***
		(0.008)	(0.009)	(0.009)
Log Likelihood		-2261.61	-2250.94	-2233.47
Observations		575	575	575

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

First, the role of geography was examined to determine if integrated circuit firms were more likely to survive where the semiconductor industry was clustered, specifically Boston, New York, Los Angeles, and Silicon Valley. Separate dummies were included for each region, with the remaining United States serving as the omitted category. The only statistically significant effect is for Silicon Valley, which is negative, indicating that firms in Silicon Valley were more likely to survive following entry into integrated circuits than firms located in other regions.

The regression presented as Model 2 combines the geography and technological background dummies in an effort to examine these two factors jointly. In addition to the variables included in Model 1, Model 2 includes dummies for each technological background, with new and unknown firms serving as the omitted category. The coefficient estimate for the transistor background variable is statistically significant negative, indicating that firms with transistor experience were more likely to survive than similar new and unknown firms. The Silicon Valley coefficient estimate remains statistically significant and negative, indicating that firms in Silicon Valley were more likely to survive following entry into integrated circuits compared to firms in other regions when controlling for location and background.

Finally, the effect of locating integrated circuit operations within a cluster (having previously operated outside of that cluster) is added to the previous specification in forming Model 3. The coefficient estimate for this variable is not statistically significant, indicating that there is no advantage to firms with respect to survival that was gained by locating operations within a transistor cluster if the firm was not previously located there. Additionally, the significance and

approximate magnitude of transistor coefficient estimate is similar to those found in the previous specification.

The results of these regressions suggest that advantageous regional effects were existent only in Silicon Valley once technological backgrounds were controlled for, and that previous transistor production was advantageous in terms of survival. Thus, we find support for Hypothesis 12 (related experience advantage) and Hypothesis 8 (regional advantage), with no support found for Hypothesis 9 (relocation advantage). Given the large proportion of firms within Silicon Valley that are spinoffs, it is possible that the regional advantage that was apparent in this analysis may actually be explained by the heritage of spinoff firms. This will be explored in a later section, at which point a fully informed discussion regarding which hypotheses are supported by the data will occur.

Market Share Attainment Analysis

The dependent variable for this analysis was constructed by collecting market share data from ICE. The publication reported the top 10 integrated circuit firms in each year by sales, which is employed in the analysis. A logit model was used to examine how several attributes contributed to the probability of attainment top market share within integrated circuit producers, with one observation for each firm included in the analysis. The dependent variable was a binary variable representing whether the firm attained top-10 market leader status by 2002 according to ICE. The same set of variables from the survival analysis is included in this market share attainment analysis, allowing us to test the same hypotheses with respect to market performance.

The sample of firms was the same as the previous analysis, and therefore descriptive statistics regarding the sample will not be repeated here. The results of the logit regression are shown below in Table 11 with positive numbers (negative) indicating a positive (negative) effect on market share attainment.

Table 12: Integrated Circuit Market Share Attainment Analysis

	N	M1	M2	M3
Location				
Boston	77	-0.65	-0.86	-0.88
		(1.10)	(1.12)	(1.12)
Los Angeles	96	N.E.	N.E.	N.E.
New York City	113	-0.99	-1.16	-1.17
		(1.10)	(1.11)	(1.12)
Silicon Valley	90	1.93***	1.77***	1.83***
		(0.59)	(0.66)	(0.67)
Relocating Firm	15			N.E.
Background				
Active Modules	43		0.37	0.35
			(1.15)	(1.15)
Diodes	23		0.47	0.41
			(1.16)	(1.17)
Electronics	128		-0.90	-0.86
			(1.12)	(1.12)
Transistors	44		1.22*	1.19
			(0.73)	(0.73)
Diversifiers	87		N.E.	N.E.
Entry Year	600	-0.12***	-0.11**	-0.12**
		(0.04)	(0.04)	(0.04)
Constant		249.9***	221.9**	238.6**
		(94.09)	(94.44)	(97.49)
Log Likelihood		-59.68	-55.66	-54.86
Observations		600	600	600

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

N.E. – Not able to be estimated.

First, the role of geography was examined to determine if integrated circuit firms were more likely to attain leading status where the semiconductor industry (and previous electronics firms) was clustered, specifically Boston, New York, Los Angeles, and Silicon Valley. Separate dummies were included for each region, with the remaining United States serving as the omitted category. The Silicon Valley coefficient estimate is statistically significant and positive, indicating that firms located in Silicon Valley were more likely to attain leading status in monolithic integrated circuits than firms located elsewhere. The coefficient estimate for year of entry is statistically significant and negative, indicating that firms that entered earlier were more likely to attain leading status, providing support for a first-mover advantage.

Next, in Model 2, the role of technological background was examined jointly with location to determine if firms with experience in related industries were more likely to attain market leadership status following their entry into integrated circuits, marked Model 2. The Silicon Valley and year of entry coefficient estimates maintain their statistical significance and magnitude in this model, and the transistor and electronics background coefficient estimates are also statistically significant. The transistor background coefficient estimate is positive and the electronics coefficient estimate is negative, indicating that firms with previous transistor experience were more likely to attain market leadership status than new firms and firms with electronics experience were less likely to do so.

Finally, the effect of locating integrated circuit operations within a cluster (having previously operated outside of that cluster) is added to the previous specification in forming Model 3. The variable could not be estimated as none of these firms attain a market leadership status. While the Silicon Valley coefficient estimate is still positive and statistically significant, it does not

appear that firms that locate their operations within a cluster, having not been located within the cluster previously, enjoyed any advantage with respect to a market leadership position. The other coefficient estimates retain their statistical significance and magnitude.

The results of the survival and market share attainment analyses indicate mixed support for some of the hypotheses. Specifically the entry analysis shows consistent support for Hypothesis 12 (background advantage), while the market share attainment analysis demonstrates support for Hypothesis 12 (background advantage) and Hypothesis 8 (regional advantage) in the case of Silicon Valley only. However, like the results from the IC survival analysis, it is entirely possible that the evidence supporting a regional effect in Silicon Valley may be explained by spinoff firms given the number of spinoff firms in the region, a possibility that will be explored later. Hypothesis 9, related to regional advantages enjoyed by movers was not supported whatsoever. As a result, it appears that background was an important attribute with respect to survival and market share attainment, but a regional advantage may explain a portion of the market share attainment story in the integrated circuit era.

A summary of the hypotheses tested and results found in this chapter are shown in the table below.

Table 13: Summary of Integrated Circuit Results

Hypothesis	Results (+/-/No support)
Hypothesis 7: Among existing firms at the time that integrated circuits are invented, firms with related electronics experience are more likely to begin producing integrated circuits than firms without electronics experience, all else held constant.	+
Hypothesis 8: Among integrated circuit producers, firms located in Boston, Los Angeles, New York and Silicon Valley are more likely to out-perform comparable firms located elsewhere, all else held constant.	+ [43]
Hypothesis 9: Among integrated circuit producers with prior electronics production, firms that locate their integrated circuit production in a different region that previous production and in the Boston, Los Angeles, New York and Silicon Valley regions are more likely to out-perform other comparable firms, all else held constant.	No Support
Hypothesis 12: Among integrated circuit entrants with previous experience, firms that had previous planar transistor production experience were more likely to perform better than other integrated circuit producers, all else held constant.	+

[43] Supported only in market share attainment analysis. May be due to spinoffs in Silicon Valley, which will be examined in Chapter 7.

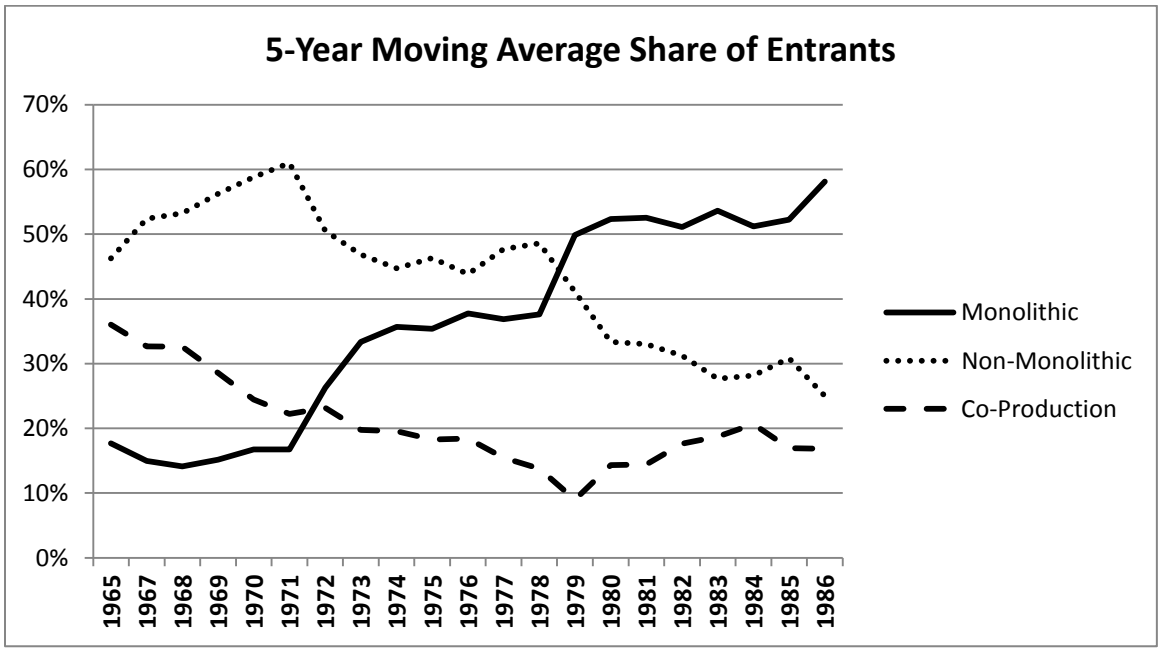
Chapter 7: Organizational Learning and Adaptation in the Semiconductor Industry

This chapter examines how firms acquire capabilities that allow them to operate at the technological frontier – both at entry and following entry. An important part of the analysis will focus on how the knowledge of spinoff firms relates to the knowledge present in their parents. As with previous analyses, it will test a set of hypotheses generated by relating general theories of firm heritage with the specific context of the semiconductor industry.

Technological Frontier at Entry

Why is production at the technological frontier so important, especially during the integrated circuit era? Monolithic integrated circuits, identified previously as the technological frontier, grew in importance in the industry over time. By 1980, they accounted for more than 90 percent of the shipped value of integrated circuits (Electronic Industries Association Marketing Services Department, 1980), making this technological frontier the centerpiece of a very large market. In response to growing demand, firms increasingly focused their activity on monolithic integrated circuits. Figure 17, below, shows that the proportion of firms producing only monolithic ICs grew from approximately 20 percent in 1965 to almost 60 percent in 1987, with declines in producers of only non-monolithic producers and producers of both types of integrated circuits.

Figure 17: 5-Year Moving Average Share of Entrants by IC Type



Source: Author's calculations from *EBG* data.

Producing at the technological frontier also had substantial implications on performance. All 37 firms that achieved top 20 market share status among integrated circuit firms produced monolithic integrated circuits. (A statistical analysis of the hypothesis regarding the performance implications of producing at the technological frontier will follow in a later section.) In an attempt to pursue this lucrative market, several firms licensed monolithic integrated circuit technology from Fairchild and Texas Instruments (Texas Instruments, 1964). But it is unclear how other firms were able to access the needed knowledge to pursue the technological frontier. Given the presence of Fairchild, one of the pioneers of monolithic integrated circuits, within Silicon Valley, it seems plausible that firms within the Valley may have been at an advantage with respect to accessing knowledge about the technological frontier.

To begin looking at organizational learning about the frontier, the likelihood of firms to produce at the technological frontier upon entry was examined first, testing the following hypothesis:

Hypothesis 15: Silicon Valley entrants in the integrated circuit industry are more likely to produce monolithic ICs at entry than entrants in the other clusters and elsewhere.

To test this, a logit regression model was used. The dependent variable is a 1-0 variable equal to 1 if the firm produced a monolithic IC in its first year of integrated circuit production and 0 otherwise. A dummy is included for each of the regional clusters to test whether production of monolithic ICs at entry is greater in Silicon Valley than in all other regions. Like before, controls for the technological background of the firm are included. The technical background dummy variables include transistors, diodes, active modules, and other electronics, in order of technical proximity to integrated circuits. Additionally, to allow for the possibility of differential effects according to the background of firms that do not have experience in closely related technologies, controls are included for diversifiers and new firms. Finally, a control for the year of entry is included to allow firms starting later to be more likely to adopt monolithic technology at entry, as its dominance became more evident over time. The sample includes 600 firms, and the background and location of the firms are shown below in table 12.

Table 14: Integrated Circuit Firms by Background and Region, 1965-1987.

Location	Active Modules	Diodes	Diversifiers	Electronics	Transistors	New	Unclassified	Total	Spinoff
Boston	9	6	14	13	5	28	2	77	8
Los Angeles	6	2	17	22	8	24	17	96	9
New York City	9	6	20	25	11	22	21	114	8
Silicon Valley	1	2	3	6	5	73	0	90	54
Other	18	7	33	62	15	53	35	223	24
Total	43	23	87	128	44	200	75	600	103

Of the 600 firms, 77 are located in Boston, 96 in Los Angeles, 113 in New York City, and 90 in Silicon Valley. The vast majority of these firms had no previous experience, with 275 new firms and 87 diversifiers, while 43 firms were active in Active Modules, 23 produced Diodes, 128 Electronics, and 44 produced Transistors. Active modules serve as the omitted category for firm background variables. The results of the regressions are shown in Table 13 below and are reported as logit coefficients, with negative (positive) numbers indicating a negative (positive) effect on producing at the technological frontier (i.e., monolithic integrated circuits) upon entry.

Table 15: Monolithic at Entry Analysis

	N	M1	M2
Location			
Boston	77	-0.23	-0.27
		(0.28)	(0.29)
Los Angeles	97	0.36	0.34
		(0.25)	(0.26)
New York City	114	0.27	0.23
		(0.24)	(0.25)
Silicon Valley	90	2.33***	2.24***
		(0.37)	(0.38)
Background			
Active Modules	43		-0.32
			(0.37)
Diodes	23		0.023
			(0.48)
Electronics	128		-0.59**
			(0.24)
Transistors	44		1.54***
			(0.39)
Diversifiers	87		-0.11
			(0.27)
Entry Year	600	0.09***	0.10***
		(0.01)	(0.01)
Constant		-178.5***	-205.5***
		(26.85)	(28.48)
Log Likelihood		-354.52	-339.93
Observations	600	600	600

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

First, the role of geography was examined, while controlling only for entry year, to determine if firms within specific regions may have had more access to knowledge regarding the technological frontier. The Silicon Valley coefficient estimate is positive and statistically significant, indicating that firms located there were more likely to produce monolithic integrated circuits upon entry. Thus they were more likely to have access to the knowledge and capabilities

required to produce at the technological frontier as compared to firms elsewhere. This result provides support for Hypothesis 15. Additionally, the coefficient estimate for the year of entry variable is also positive and statistically significant, indicating that firms were more likely to produce monolithic integrated circuits the later that they began producing integrated circuits. This seems reasonable given the growing importance of monolithic integrated circuit production over time, as noted above.

Given the historical account on the evolution of the technology, one might also expect that firms with previous technological backgrounds, specifically in transistor production, might have been more likely to produce monolithic integrated circuits upon entry than other firms. To analyze this, technological background variables were added to the variables previously included. The results, shown in Model 2, demonstrate that the Silicon Valley effect remains statistically significant, and that the coefficient estimate for the transistor variable is positive and statistically significant. This result confirms that firms with previous transistor production were more likely to produce monolithic integrated circuits at entry than other similar firms. This lends support to the notion that firms pursue specific technologies based on their previous background. However, the inclusion of the background variables does not eliminate the significance of the Silicon Valley effect, indicating that both location and background played a role with regard to a firm's production at the frontier upon entry. These findings raise raises the question as to how firms in that specific region were more likely to produce at the technological frontier. This may have been due to knowledge spillovers within the Silicon Valley region or because of knowledge dissemination through spinoff firms in the region. This topic will be examined further in future sections, specifically through examining the overlap in production between spinoffs and parents.

Adoption of Technological Frontier Following Entry

While some firms began producing at the technological frontier once they entered ICs, others did not. Since being at the frontier was critical for performance, it is important to investigate what allowed firms that did not start as producers of monolithic integrated circuits to move to the technological frontier following entry. We know that one of the leading firms with respect to monolithic integrated circuits, Fairchild, was located in Silicon Valley. Agglomeration theory predicts that firms located near Fairchild and other monolithic integrated circuit producers may be able to gain access to technological knowledge through knowledge spillovers. As a result, it is expected that firms in regions with a strong presence of monolithic integrated circuit producers adopt monolithic integrated circuit technology at a faster pace than firms located elsewhere, testing this hypothesis:

Hypothesis 11: The rate at which integrated circuit producers adopt monolithic ICs at any given moment is greater for producers located in clusters of monolithic integrated circuit producers than other similar firms located outside of the these clusters.

For this analysis, firms that did not use monolithic technology when they entered integrated circuits are considered, looking at their subsequent decision to adopt monolithic integrated circuits. A Cox Proportional Hazard Model is estimated where, similarly to the previous regression, four regional dummies are also included, as well as background dummies and a time of entry variable to allow for heterogeneity across firm characteristics, regions and time. Firms located in clusters are expected to be more likely to adopt monolithic technology earlier due to the possible presence of knowledge spillovers. Firms with related technological background would also be more likely to begin producing monolithic integrated circuits earlier due to the

similarity in knowledge base. The sample included 300 integrated circuit producers that did not product monolithic ICs upon entry, of which 36 – slightly more than 10 percent - adopted monolithic integrated circuits following entry. This proportion is striking, given that almost half of all integrated circuit producers produced monolithic integrated circuits upon entry. The inability of the vast majority of the remaining 300 firms to produce monolithic integrated circuits speaks volumes to either a lack of motivation to do so, which doesn't seem likely, or an inability to access the required knowledge and skills to do so.

Of the 300 firms included in the sample, 50 were located in Boston, 51 in Los Angeles, 60 in New York City, and 10 in Silicon Valley. Only firms that entered prior to 1987 were included in this sample because adoption after entry could not be observed for firms that started production in the last year of our data. Like before, some of these firms had previous experience in the industry, with 28 making Active Modules, 12 producing Diodes, 83 Electronics, and 12 Transistors. The results of the regressions are shown below in Table 14 and are reported as hazard coefficients, with negative (positive) numbers indicating a negative (positive) effect on adopting the technological frontier following entry.

Table 16: Adoption of Monolithic ICs Following Entry

	N	M1	M2
Location			
Boston	50	-0.69	-0.74
		(0.56)	(0.62)
Los Angeles	51	-0.19	-0.048
		(0.51)	(0.54)
New York City	60	0.24	0.15
		(0.41)	(0.42)
Silicon Valley	10	-0.19	-0.32
		(0.76)	(0.87)
Background			
Active Modules	28		0.52
			(0.60)
Diodes	12		0.59
			(0.76)
Electronics	83		0.20
			(0.45)
Transistors	12		1.26**
			(0.60)
Diversifiers	48		-0.22
			(0.65)
Entry Year	300	-0.03	-0.01
		(0.03)	(0.03)
Observations		300	300
Log Likelihood		-149.45	-146.80

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

First, in Model 1, the role of geography was examined, while controlling for entry year, to determine if firms within specific regions may have had more access to knowledge regarding the technological frontier. No coefficient estimates are statistically significant, indicating that firms located throughout the United States were just as likely to begin producing monolithic integrated circuits following entry. This means there were no regional effects for learning about the technological frontier over time, which is in direct contrast to the expectation that firms located

within Silicon Valley would be able to utilize knowledge spillovers from Fairchild and other spinoffs in the region as well as hiring and other interactions to gain access to monolithic integrated circuit knowledge. No support is found for Hypothesis 11. Additionally, the coefficient estimate for the year of entry variable is not statistically significant; indicating that time of entry was not a factor with regard to whether the firm was able to adopt monolithic integrated circuit production following entry.

In Model 2, the coefficient estimate for the transistors variable is positive and statistically significant, indicating that firms with previous transistor production were more likely to adopt monolithic integrated circuits following entry than other similar firms. This finding lends support to the notion that firms are able to pursue specific technologies based on their previous background, and it appears that only background played a role with regard to a firm's adoption of the technological frontier following entry.

While the ability of firms to produce at the technological frontier upon entry appears to be related to both technological background and the region in which the firm operates, only prior technology plays a role with respect to adoption of the technological frontier following entry. This implies that if regional effects are in play in Silicon Valley, these regional effects are only effective at the time of entry, limiting spillovers or hiring effects to those that are possible when the firm is founded. One possible explanation for this regional effect would be a large number of spinoffs that are founded by individuals that previously worked at firms that had produced monolithic integrated circuits, and the role of such a situation in the pursuit of monolithic integrated circuits upon entry will be explored later.

Spinoff Learning through Experience Gained at Parent

The final section of this chapter examines the role of knowledge transfer through the spinoff process. Previous literature has shown that spinoffs are more likely to out-perform other entrants (Klepper, 2009b). But we know much less about what allows spinoffs to achieve such success. Some have suggested that spinoffs learn from their parent firms (Franco & Filson, 2006; Klepper & Sleeper, 2005; Moore & Davis, 2004). However, it could also be true that spinoffs are simply the firms most able to take advantage of knowledge spillovers that exist within a given region. Both conduits for knowledge spillovers are examined in an effort to determine which one(s) are in play in the integrated circuit era.

The first hypothesis to be examined relates to the dissemination of knowledge required for producing monolithic integrated circuits through the spinoff mechanism:

Hypothesis 13: Firms founded by individuals with previous experience producing monolithic ICs are more likely to produce monolithic ICs than firms with other backgrounds, all else held constant.

To test this hypothesis, data on production of the spinoff and parent firm within three years of the spinoff's initial production were gathered. The previous analysis on the propensity of firms to produce products at the technological frontier upon entry was then extended using these data. The regression is similar to what was found in Table 13, to which a binary dummy for spinoffs whose parent firms produced monolithic integrated circuits at the time that the spinoff was founded was added. The results are presented in Table 15:

Table 17: Expanded Monolithic at Entry Analysis

	N	M1	M2	M3
Location				
Boston	77	-0.23	-0.27	-0.26
		(0.28)	(0.29)	(0.29)
Los Angeles	97	0.36	0.34	0.34
		(0.25)	(0.26)	(0.27)
New York City	114	0.27	0.23	0.26
		(0.24)	(0.25)	(0.25)
Silicon Valley	90	2.33***	2.24***	1.79***
		(0.37)	(0.38)	(0.39)
Background				
Active Modules	43		-0.32	-0.058
			(0.37)	(0.38)
Diodes	23		0.023	0.32
			(0.48)	(0.49)
Electronics	128		-0.59**	-0.31
			(0.24)	(0.25)
Transistors	44		1.54***	1.84***
			(0.39)	(0.40)
Diversifiers	87		-0.11	0.16
			(0.27)	(0.28)
Spinoff of Monolithic Parent	74			1.63***
				(0.40)
Entry Year	600	0.09***	0.10***	0.10***
		(0.01)	(0.01)	(0.01)
Constant		-178.5***	-205.5***	-212.9***
		(26.85)	(28.48)	(29.15)
Log Likelihood		-354.52	-339.93	-330.64
Observations	600	600	600	600

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

The previous results are labeled Models 1 and 2 for the sake of comparison, with Model 3 serving as the new specification. The coefficient estimate for the variable representing spinoff firms that may have been able to access monolithic integrated circuit knowledge is positive and statistically significant at the 1 percent level, indicating that these firms were more likely to

produce monolithic integrated circuits upon entry as compared to other similar firms. This result provides support for Hypothesis 13 and the notion that product knowledge was passed from parent to spinoff. Although inclusion of this variable decreases the magnitude of the Silicon Valley coefficient estimate by approximately twenty percent, the statistical significance of the variable remains. The remaining significance of the Silicon Valley coefficient estimate may be due to imperfect identification of the background of IC producers, driven by two issues. First, there are 75 firms that are unclassified following the extensive effort to classify firms, some of which could be spinoff firms or of another classification. Second, with respect to those firms that are classified as spinoffs, some of the spinoff parents are not included in the EBG and thus there is no detailed data regarding production. For any of these firms that fall into the second category, the dummy regarding parent monolithic production is set to zero. These two issues together may result in a downward bias of the coefficient estimate, causing the Silicon Valley coefficient estimate to retain its significance.

If the presence of a parent firm in monolithic integrated circuits provides a strong explanation of the ability of its spinoffs to begin with equivalent capabilities, one may assume that the founders of spinoffs are learning about these technologies in the parent firms. If this is the case, it seems reasonable to think that this transfer of knowledge from parent to spinoff was not specific to monolithic integrated circuit product but rather a common characteristic of spinoffs in all technologies. Therefore, all integrated circuit production can be examined to test such an overlap in production:

Hypothesis 14: Among spinoff firms, the probability that a spinoff firm produces a given product will be greater if the spinoff's parent previously produced the product.

Using the spinoff and founder information mentioned in Chapter 4, parent firms were identified for 108 of the entrants in the industry, with 8 in Boston, 9 in Los Angeles, 9 in New York, 54 in Silicon Valley and 28 located outside the main industry clusters. The important presence of Silicon Valley in this sample reflects the large amount of research that has focused on the origin of firms located in this region. The analysis included all spinoff firms with parents that had production data in the *EBG* (99).

To analyze the technology overlap between parents and spinoffs, the IC products in the eight broad categories were considered. If a parent firm produces a specific type of IC, we expect its spinoff to be more likely to produce that IC type when compared with the average spinoff. Thus, up to eight observations were created for each spinoff, corresponding to the broad IC types actively being produced by firms in the industry at the time the spinoff entered. Table 16, below, describes the eight broad product categories.

Table 18: Broad Integrated Circuit Product Categories

Broad Product Category	First Production Year	Last Production Year	Peak Production Year	Firms Producing in Peak Year
Film	1965	1987	1972	116
Hybrid	1965	1987	1969	145
Monolithic Custom	1970	1987	1987	67
Monolithic Linear	1970	1987	1987	97
Monolithic Logic	1970	1987	1975	62
Monolithic Memory	1970	1987	1987	45
Monolithic Microprocessor	1976	1987	1987	52
Monolithic Other	1965	1987	1987	84

For each product type and spinoff, a production dummy was generated, which equals 1 if the parent produced that given product within one year before or after the spinoff's entry and 0 otherwise. Similarly, the corresponding dependent variable was set equal to 1 if the spinoff produced the product type within one year of its entry and 0 otherwise.

A probit model was then estimated where the coefficient estimate on the parent production dummy is expected to be greater than 0, reflecting an increased likelihood that a spinoff firm produced a given product if its parent firm produced it. In the estimation, we also include dummies for the eight product types, year of entry, and a Silicon Valley region dummy to control for overlaps in spinoff and parent production due to any of the particular factors of the region. The Silicon Valley dummy was included due to the large proportion of spinoff firms that are founded in Silicon Valley. Given that the region has been heavily studied, there is a large amount of firm heritage data available for firms there, raising a concern that the more short-lived, lower performing spinoffs may have been identified in Silicon Valley as compared to other regions. This dummy served to control for any difference in product knowledge inheritance due to this perceived difference in spinoff identification information. Standard errors were computed by clustering observations for each parent firm. The coefficient estimates of the probit regressions are reported below in Table 17, with year and product coefficient estimates suppressed.

Table 19: Likelihood (Probit) that Spinoff Produces a Product Given that Parent Produced It - Aggregate Product Categories

	N	M1	M2	M3	M4
Parent Production	392	0.87***	0.87***	0.81***	0.84***
		(0.14)	(0.14)	(0.17)	(0.17)
Product Dummies	523		Included	Included	Included
Entry Year Dummies	523			Included	Included
Silicon Valley	361				0.14
					(0.28)
Constant		-0.93***	-0.80***	-1.16***	-1.34***
		(0.13)	(0.18)	(0.31)	(0.35)
Log Likelihood		-332.16	-320.44	-279.64	-279.12
Observations	523	523	523	514	514

Products indicate dummies included to control for each of the 8 product types. Years indicate dummies to control for spinoff entry year.
Robust Standard Errors in Parentheses; *** Significant at the .01 level; ** at .05 level; * at .10 level
Sample includes all spinoff firms with an identified parent that is present and actively producing in the *EBG* within three years of the spinoff's entry.

Model 1 includes only the parent production variable. The parent coefficient estimate is positive and significant at the .01 level, implying that spinoff firms were more likely to produce a given product if their parent also produced it. The estimate does not change much and remains significant at the .01 level with the inclusion of product (M2), year of entry (M3), and Silicon Valley (M4) dummies. The Silicon Valley coefficient estimate is not statistically significant, suggesting that there is not much of a difference in the ability of spinoffs to learn from their parents inside and outside that region.

To test the hypothesis at a narrower level of production, the eight product categories were further disaggregated into 50 types that were either technically different or served different markets.[44] Up to 50 observations for each spinoff were created corresponding to the narrow IC types that were actively being produced by firms in the industry at the time the spinoff entered. After the disaggregation of product categories, a total of 1,341 firm-product-year observations remain. Standard errors were again computed by clustering observations at the parent firm-product level. Coefficient estimates for the model are reported below in Table 18, with the year and product coefficient estimates suppressed.

[44] For example, various types of logic circuits such as Resistor-Transistor Logic (RTL) and Transistor-Transistor Logic (TTL) were separated in this analysis whereas Microprocessors and Microcontrollers could not be separated as they were actually reported as a single category for several years.

Table 20: Likelihood (Probit) that Spinoff Produces a Product Given that Parent Produced It - Disaggregated Product Categories

	N	M1	M2	M3	M4
Parent Production	654	0.80***	0.68***	0.57***	0.58***
		(0.097)	(0.10)	(0.11)	(0.11)
Product Dummies	1,341		Included	Included	Included
Entry Year Dummies	1,341			Included	Included
Silicon Valley	1,104				-0.074
					(0.12)
Constant		-1.48***	-1.94***	-2.14***	-2.14***
		(0.075)	(0.17)	(0.35)	(0.35)
Log Likelihood		-537.31	-478.55	-416.91	-416.72
Observations	1,341	1,341	1,332	1,332	1,332

Products indicate dummies included to control for each of the 8 product types. Years indicate dummies to control for spinoff entry year.
Robust Standard Errors in Parentheses; *** Significant at the .01 level; ** at .05 level; * at .10 level
Sample includes all spinoff firms with an identified parent that is present and actively producing in the *EBG* within three years of the spinoff's entry.

The coefficient estimates for the parent production variable are again positive and significant at the .01 level regardless of whether controls for year, region, and product are included. Again, the coefficient estimate of the Silicon Valley dummy is not statistically significant, indicating that there was no regional effect with respect to spinoffs entering specific products.

These results suggest that, as stated in Hypothesis 14, spinoffs are likely to produce a subset of the IC products their parent was producing. To convey the extent of the overlap, Table 19

presents the broad IC categories produced by parents and their spinoffs and reports the average percentage of parent categories also produced by their spinoffs.[45]

Table 21: Share of Integrated Circuit product categories produced by parents and spinoffs

Categories produced by:	Parent (mean)	Spinoff (mean)	Parent Only (share)	Spinoff Only (share)	Both Firms (share)
Broad Technical areas (8)	5.0	2.7	39%	4%	36%
Narrow Technical areas (50)	9.9	4.1	37%	4%	12%

The overlap between spinoff and parent production is striking, as is the fact that spinoffs tend to systematically produce fewer product types than their parent. In fact, only 4% of the spinoffs produced product types that their parent did not produce at their time of entry at both the broad and narrow product levels, a remarkably small amount. Moreover, spinoffs tended to produce only a subset of the products their parents did. At the broad product level they produced 36% of the product types their parents did; but only 12% of the product types at the narrow level. In contrast, nearly 40% of the parents were active in IC product categories that their spinoffs were not active in - at either the broad or the narrow level.

[45] Specifically, the table lists the share of products produced by each firm in a 3-year window given what was available at the time, averaged across all firms. For parent firms, the window consists of one year prior to spinoff entry, the year of spinoff entry, and the year following spinoff entry. For spinoff firms, the window includes the year of entry as well as the two years following entry.

Spinoff Performance and Influence of Technological Frontier Production

We have seen strong evidence that spinoff firms gain knowledge from the founder's previous work experience at parent firms, manifested through the types of products that spinoffs are able to produce at the time of founding. While previous work has shown that spinoffs are able to out-perform other firms because of the knowledge that they are able to gain from previous employment (Klepper, 2007a), this has not yet been tested with the semiconductor industry data. Additionally, we have seen evidence suggesting that operating at the technological frontier was important in both the transistor and integrated circuit eras. However this also has not been analyzed with the data. This section will test hypotheses related to the performance of spinoff firms as well as hypotheses related to the performance of firms producing at the technological frontier.

Transistor Analysis

First, the role of production at the technological frontier with respect to firm performance will be examined. From the discussion of technological development during the transistor era in Chapter, 3, it may have been evident that the importance of the use of silicon in the production of transistors increased over time. In fact, it was only through the use of silicon in the production of transistors that oxide masking was able to occur, a key manufacturing step required for mesa and planar transistors (Riordan & Hoddeson, 1997, p. 222). In a similar fashion to the integrated circuit era, all 12 of the leading transistor firms produced silicon transistors, leading one to wonder what role producing at the technological frontier had on firm performance. While no specific hypothesis has been developed with respect to the transistor era, it seems reasonable to

adapt the hypothesis regarding the production of monolithic integrated circuits to silicon transistor production during the transistor era:

Hypothesis 17: Firms that produced silicon transistors upon entry were more likely to out-perform other firms that did not produce silicon transistors upon entry, all else held constant.

The results above indicated that spinoffs from high quality integrated circuit parents were more likely to out-perform other new firms, presumably because of the complex, tacit knowledge passed from parent to spinoff. However, we do not expect the same to be true during the transistor era. While the knowledge required for producing monolithic integrated circuits was complex, tacit, and centered within two key firms (Fairchild and Texas Instruments), which seems to have given an advantage to spinoff firms, the situation was very different in the transistor era. As detailed in Chapter 3, Bell Labs was very open in disseminating information regarding transistor technology to interested parties. The openness of Bell, as well as the distributed nature of transistor innovations may have mitigated any advantages that spinoffs would have by being able to utilize knowledge and experience gained during previous employment. As a result, we may see performance of spinoff firms be comparable to other new entrants, as described in the following hypothesis:

Hypothesis 6: Among transistor producers, spinoff firms should perform at a comparable level to other firms, all else held constant.

To analyze both of these hypotheses, the survival and market share attainment analyses presented in Chapter 5 are built upon.

First, a dummy for all spinoffs whose parent firm had achieved top market share was included in order to examine whether spinoffs from high performing parents out-performed other firms, as specified in Hypothesis 6. This factor was added to the survival and market share attainment analyses (shown previously in Tables 6 and 7, respectively), and results of the survival and market share attainment analyses are marked as Model 4 in Tables 20 and 21 below, respectively. The coefficient estimate for the variable representing spinoffs from high quality parents is negative and statistically significant in the survival analysis, but is not statistically significant in the market share attainment analysis, indicating that a survival advantage (but no market share attainment advantage) existed for spinoff firms of high quality parents when these firms were compared to other firms, which provides partial support for Hypothesis 6 (transistor spinoffs are comparable).

The influence of producing at the technological frontier upon entry on firm performance was analyzed next. This factor was added to the survival and market share analyses by creating a dummy for firms that produced silicon transistors during the year of initial transistor production, and this dummy was added to each of the models. Results of the survival and market share attainment analyses are marked as Model 5 in Tables 20 and 21 below, respectively. The coefficient estimates for this variable are not statistically significant, indicating that there was no survival or market share attainment advantage to producing at the technological frontier upon entry.

When all variables of interest are included, tube producers, producers of silicon transistors and early entrants appear to be the firms most likely to perform well following entry into transistor production. It should be noted that the Silicon Valley coefficient estimate retains its significance

in the market share attainment analysis. However, only one firm – Fairchild – drives this effect. Fairchild is the only firm in Silicon Valley that achieves market share leadership status, and excluding this firm from the analysis would eliminate the significance of the coefficient estimate, which would not be estimated because of perfect prediction of failures. To recap, firms that had related experience, that were at the technological frontier, and that were early movers were most likely to perform the best. Given that the firms that enjoyed entry and performance advantages based on their background (specifically tube producers) were located primarily within the existing electronics clusters, the transition from tubes to transistors served to reinforce the existing electronics clusters. The same geographic inertia would not play out during the transition to integrated circuits as the emergence of spinoff firms moved the industry to Silicon Valley. In fact, these results contrast with the significant advantage that spinoffs enjoyed during the later integrated circuit era.

Table 22: Expanded Transistor Survival Analysis

	N	M1	M2	M3	M4	M5
Location						
Boston	13	-0.21	-0.19	-0.22	-0.23	-0.23
		(0.35)	(0.36)	(0.36)	(0.37)	(0.37)
Los Angeles	13	-0.16	-0.24	-0.27	-0.04	-0.04
		(0.34)	(0.35)	(0.36)	(0.36)	(0.36)
New York City	47	0.05	0.14	0.12	0.01	0.01
		(0.23)	(0.24)	(0.24)	(0.24)	(0.24)
Silicon Valley	7	-0.92*	-0.91*	-0.93*	-0.52	-0.52
		(0.49)	(0.50)	(0.50)	(0.52)	(0.52)
Relocating Firm	6			-0.39	-0.41	-0.41
				(0.49)	(0.49)	(0.49)
Background						
Amplifiers	6		-0.29	-0.21	-0.34	-0.34
			(0.49)	(0.50)	(0.50)	(0.50)
Diversifiers	18		0.05	0.05	-0.15	-0.15
			(0.30)	(0.30)	(0.31)	(0.31)
Electronics	24		0.169	0.22	0.10	0.10
			(0.26)	(0.27)	(0.27)	(0.27)
Tubes	9		-0.82*	-0.82*	-0.88*	-0.88*
			(0.46)	(0.46)	(0.46)	(0.46)
Other Electronics	8		-0.51	-0.45	-0.55	-0.55
			(0.44)	(0.45)	(0.45)	(0.45)
Spinoff of Top Transistor Firm	10				-1.06**	-1.06**
					(0.49)	(0.49)
Silicon Transistor Production at Entry	52					0.01
						(0.24)
Entry Year	118	0.07**	0.04	0.04	0.06*	0.06
		(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Log Likelihood		-426.53	-423.48	-423.12	-420.47	-420.46
Observations	118	118	118	118	118	118

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 23: Expanded Transistor Market Share Attainment Analysis

	N	M1	M2	M3	M4	M5
Location						
Boston	13	-0.41	-0.51	-0.30	-0.25	-0.25
		(1.01)	(1.27)	(1.30)	(1.32)	(1.36)
Los Angeles	13	0.80	1.26	1.49	1.31	1.54
		(1.06)	(1.17)	(1.22)	(1.26)	(1.30)
New York City	47	-1.58	-3.24*	-3.27	-3.33	-3.93
		(1.20)	(1.89)	(2.03)	(2.12)	(2.54)
Silicon Valley	7	2.48	3.12*	3.45*	3.16*	3.40*
		(1.61)	(1.76)	(1.86)	(1.91)	(1.95)
Relocating Firm	6			1.67	0.91	0.86
				(1.72)	(2.00)	(2.00)
Background						
Amplifiers	6		1.14	0.92	1.10	0.91
			(1.45)	(1.43)	(1.49)	(1.52)
Diversifiers	18		-0.28	-0.27	0.03	-0.01
			(1.13)	(1.14)	(1.23)	(1.23)
Electronics	24		0.08	-0.08	0.20	-0.03
			(1.44)	(1.47)	(1.54)	(1.60)
Tubes	9		3.22*	3.41*	3.69*	4.39*
			(1.81)	(1.94)	(2.08)	(2.52)
Other Electronics	8		0.41	-0.26	-0.07	0.53
			(1.47)	(1.73)	(1.77)	(1.79)
Spinoff of Top Transistor Firm	9				1.38	1.27
					(1.78)	(1.75)
Silicon Transistor Production at Entry	52					2.22
						(1.71)
Entry Year	118	-0.62***	-0.65***	-0.69***	-0.73***	-1.02***
		(0.20)	(0.22)	(0.24)	(0.25)	(0.38)
Constant		1,218***	1,271***	1,358***	3,505**	3,454**
		(390.7)	(441.2)	(477.9)	(1,544)	(1,539)
Log Likelihood		-24.27	-21.65	-21.20	-20.90	-20.01
Observations		118	118	118	118	118

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Integrated Circuit Analysis

To compare the two eras, similar analyses will be performed for the integrated circuit era. First, the importance of producing at the technological frontier will be examined. As discussed earlier, monolithic integrated circuits accounted for a vast majority of the value of integrated circuits shipped by 1980, and every top 20 integrated circuit producer was active in monolithic integrated circuits upon entry. Such an advantage can be analyzed by testing the following hypothesis:

Hypothesis 10: Firms that were producing monolithic integrated circuits at entry were more likely to out-perform other firms that did not produce monolithic integrated circuits, all else held constant.

As mentioned previously, the literature has also shown that spinoffs are able to out-perform comparable firms. It is believed that the mechanism that provides these firms with an advantage with respect to other entrants is the knowledge that they are able to gain from previous employment, which was demonstrated in a previous section of this chapter. Given the strong connection between products, it seems reasonable to expect the same advantage in terms of performance, with the specific expectation defined in the following hypothesis:

Hypothesis 16: Among integrated circuit producers, spinoffs that have a high quality parent firm are likely to out-perform other firms, all else held constant.

To analyze both of these hypotheses, and to examine the role of mechanisms that play a role at entry versus effects over time within regions, these hypotheses will be examined below by building on the survival and market share attainment analyses that were performed in Chapter 5.

First, a dummy for all spinoffs whose parent firm had achieved top 10 market share was included in order to examine whether spinoffs from high performing parents out-performed other firms, as specified in Hypothesis 16. This factor was added to the survival and market share attainment analyses (shown previously in Tables 10 and 11, respectively), and results of the survival and market share attainment analyses are marked as Model 4 in Tables 22 and 23 below, respectively. The coefficient estimate for the variable representing spinoffs from high quality parents is statistically significant in both analyses, with the signs of the coefficient estimates in each analysis indicating a survival and market share attainment advantage for spinoff firms of high quality parents as compared to other firms. Additionally, the magnitude of the Silicon Valley coefficient estimate (which remains marginally significant in the market share attainment analysis) is reduced by the inclusion of this variable. Given that the measure of having a high quality parent included a mix of top performing and less successful spinoffs, the remaining significance of the regional effect is not entirely surprising. Without having detailed data on the quality of spinoff firms separate from parent information, however, this is the best measure that can be used to test Hypothesis 16. These findings provide support for Hypothesis 16 (spinoff advantage) and the idea that spinoffs are able to capitalize on the knowledge that they gain from their parent firm, specifically high quality parent firms, providing an advantage with respect to performance.

Building on the previous analysis, the influence of producing at the technological frontier at entry on firm performance was analyzed. This factor was added to the survival and market share attainment analyses by creating a dummy for firms that produced monolithic integrated circuits during the first year of production, and this dummy was added to each of the models. Results of

the survival and market share attainment analyses are marked as Model 5 in Tables 22 and 23 below, respectively. The coefficient estimate for this variable could not be estimated for the market share attainment regression because all firms that achieved market leadership produced monolithic integrated circuits, which is consistent with our hypothesis. But the coefficient estimate for the survival regression is negative and statistically significant, indicating that there was indeed a survival advantage to producing at the technological frontier. Given the survival effect of this variable, it may be serving as a proxy for high quality firms, which combined with the inclusion of the high quality parent dummy for spinoff firms eliminates the significance of the Silicon Valley coefficient estimate for market share attainment.

Table 24: Expanded Integrated Circuit Survival Analysis

	N	M1	M2	M3	M4	M5
Location						
Boston	75	-0.26	-0.23	-0.23	-0.2	-0.23
		(0.17)	(0.17)	(0.17)	(0.17)	(0.17)
Los Angeles	94	0.09	0.05	0.06	0.04	0.08
		(0.14)	(0.14)	(0.14)	(0.14)	(0.14)
New York City	111	0.09	0.12	0.12	0.13	0.14
		(0.13)	(0.13)	(0.13)	(0.13)	(0.13)
Silicon Valley	80	-0.34**	-0.32*	-0.32*	-0.10	0.001
		(0.17)	(0.17)	(0.18)	(0.18)	(0.19)
Relocating Firm	13			0.02	0.01	-0.047
				(0.32)	(0.32)	(0.33)
Background						
Active Modules	42		0.10	0.10	0.021	0.023
			(0.19)	(0.19)	(0.19)	(0.19)
Diodes	21		-0.48	-0.48	-0.60**	-0.57*
			(0.30)	(0.30)	(0.30)	(0.30)
Electronics	122		0.03	0.03	-0.05	-0.05
			(0.13)	(0.13)	(0.13)	(0.13)
Transistors	43		-0.79***	-0.79***	-0.91***	-0.81***
			(0.22)	(0.22)	(0.22)	(0.23)
Diversifiers	84		0.11	0.11	0.018	0.03
			(0.15)	(0.15)	(0.15)	(0.15)
Spinoff of Top 10 Parent	47				-0.89***	-0.81***
					(0.26)	(0.26)
Monolithic Production at Entry	279					-0.25**
						(0.11)
Entry Year	575	-0.03***	-0.03***	-0.03***	-0.03***	-0.03***
		(0.008)	(0.009)	(0.009)	(0.009)	(0.009)
Log Likelihood		-2261.61	-2250.94	-2250.94	-2243.92	-2241.41
Observations		575	575	575	575	575

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 25: Expanded Integrated Circuit Market Share Attainment Analysis

	N	M1	M2	M3	M4	M5
Location						
Boston	77	-0.65	-0.86	-0.88	-0.96	-0.91
		(1.10)	(1.12)	(1.12)	(1.13)	(1.19)
Los Angeles	96	N.E.	N.E.	N.E.	N.E.	N.E.
New York City	113	-0.99	-1.16	-1.17	-1.18	-1.37
		(1.10)	(1.11)	(1.12)	(1.13)	(1.17)
Silicon Valley	90	1.93***	1.77***	1.83***	1.27*	0.85
		(0.59)	(0.66)	(0.67)	(0.74)	(0.81)
Relocating Firm	15			N.E.	N.E.	N.E.
Background						
Active Modules	43		0.37	0.35	1.21	2.34*
			(1.15)	(1.15)	(1.29)	(1.41)
Diodes	23		0.47	0.41	1.47	1.93
			(1.16)	(1.17)	(1.29)	(1.42)
Electronics	128		-0.90	-0.86	0.03	0.4
			(1.12)	(1.12)	(1.27)	(1.38)
Transistors	44		1.22*	1.19	2.17**	1.53
			(0.73)	(0.73)	(0.94)	(1.02)
Diversifiers	87		N.E.	N.E.	N.E.	N.E.
Spinoff of a Top 10 Parent	51				2.35***	1.95**
					(0.87)	(0.87)
Monolithic Production at Entry	300					N.E.
Entry Year	600	-0.12***	-0.11**	-0.12**	-0.12**	-0.15***
		(0.04)	(0.04)	(0.04)	(0.05)	(0.05)
Constant		249.9***	221.9**	238.6**	249.5**	307.9***
		(94.09)	(94.44)	(97.49)	(105.2)	(105.9)
Log Likelihood		-59.68	-55.66	-54.86	-50.43	-43.20
Observations		600	600	600	600	600

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

The results of the analyses in this chapter support many of the hypotheses that were examined. A summary of the hypotheses tested and results found in this chapter are shown in the table below.

Table 26: Summary of Organizational Learning and Adaptation Results

Hypothesis	Results (+/-/No support)
Hypothesis 2: Among transistor producers, firms with related electronics experience, specifically tube manufacturers, may be more likely to perform better than other firms, all else held constant.	+
Hypothesis 3: Among transistor producers, firms that begin producing transistors earlier are more likely to perform better than other firms, all else held constant.	+
Hypothesis 4: Among transistor producers, firms located in Boston, Los Angeles and New York are more likely to perform better than firms located elsewhere, all else held constant.	No Support
Hypothesis 5: Among transistor producers with prior electronics production, firms that locate their transistor production in a different region that previous production and in the Boston, Los Angeles, and New York regions are more likely to out-perform other comparable firms, all else held constant.	No Support
Hypothesis 6: Among transistor producers, spinoff firms should perform in a similar to fashion to other comparable firms, all else held constant.	+ [46]
Hypothesis 8: Among integrated circuit producers, firms located in Boston, Los Angeles, New York and Silicon Valley are more likely to out-perform comparable firms located elsewhere, all else held constant.	No Support
Hypothesis 9: Among integrated circuit producers with prior electronics production, firms that locate their integrated circuit production in a different region that previous production and in the Boston, Los Angeles, New York and Silicon Valley regions are more likely to out-perform other comparable firms, all else held constant.	No Support
Hypothesis 10: Firms that were producing monolithic integrated circuits were more likely to out-perform other firms that did not produce monolithic integrated circuits, all else held constant.	+

[46] Hypothesis only supported in the market share attainment analysis.

Hypothesis 11: The rate at which integrated circuit producers adopt monolithic ICs at any given moment is greater for producers located in clusters of monolithic integrated circuit producers than other similar firms located outside of the these clusters.	+ [47]
Hypothesis 12: Among integrated circuit entrants with previous experience, firms that had previous planar transistor production experience were more likely to perform better than other integrated circuit producers, all else held constant.	+
Hypothesis 13: Firms founded by individuals with previous experience producing monolithic ICs are more likely to produce monolithic ICs than firms with other backgrounds, all else held constant.	+
Hypothesis 14: Among spinoff firms, the probability that a spinoff firms produces a given product will be greater if the spinoff's parent previously produced the product.	+
Hypothesis 15: Silicon Valley entrants in the IC industry are more likely to produce monolithic ICs at entry than entrants in the other clusters and elsewhere.	+
Hypothesis 16: Among integrated circuit producers, spinoffs that have a high quality parent firm are likely to out-perform other firms, all else held constant.	+

Firms located in Silicon Valley, as well as those that entered early and had previous transistor experience were more likely to produce at the technological frontier of integrated circuits upon entry. However, after entry, only transistor experience seemed to play a role in allowing firms to adopt products at the technological frontier, with no evident effects of knowledge spillovers. The appearance of a regional effect in Silicon Valley with respect to the production at the technological frontier upon entry may be tied to the large proportion of spinoffs there, as the next two analyses show evidence of a link between the production of spinoff and parent firms.

It is evident that spinoffs rely on knowledge gained from parents regarding technology and markets in order to manufacture products. This knowledge dissemination occurs not only with

[47] Support shown may be due to incomplete identification of spinoff firms and production of spinoff parents.

the technological frontier knowledge, but in general about the broad sets of activities that parents perform. The use of this knowledge allowed spinoffs to enjoy a distinct advantage in terms of firm performance in the integrated circuit era. These findings support the notion that spinoffs were able to access technology and market knowledge which influenced the spinoff's product portfolio, as well as management knowledge as suggested by Moore and Davis (2004), which may have influenced the spinoff's performance. Spinoff firms in the transistor era did not enjoy a similar advantage most likely because of the accessibility of transistor knowledge from Bell Labs and other leading industry sources such as RCA. The complex and tacit nature of the knowledge required to produce integrated circuits, however, made this knowledge dissemination between parent and spinoff advantageous for spinoff firms, especially if this knowledge was coming from a top performing firm.

For both the transistor and integrated circuit era, we find that firms from related fields (transistors and tubes), those at the technological frontier, as well as early movers, enjoyed advantages for both survival and performance. Yet, the increasingly complex and tacit nature of knowledge that was associated with integrated circuits appears to have allowed spinoffs to out-perform other firms, turning what was previously not a strong electronics cluster (Silicon Valley) to the main cluster of the industry.

Chapter 8: Conclusions & Policy Implications

In an effort to understand how new industries emerge, the semiconductor industry was examined to gain understanding as to specifically what explains the establishment and development of new industries. Examining during two distinct eras (transistors and integrated circuits) of this industry has uncovered similarities and differences with regard to firm entry and performance between the two eras, but more importantly exposed a shift in the mechanism that better explains firm performance across these two eras.

During both eras of this new industry, related experience was a very significant factor in explaining firm entry in this emerging sector. This observation supports the main tenets of heritage theory, which predicts that firms with related experience are the most likely among existing producers to being producing new, related products that emerge in a new industry. During the transistor era, previous experience in vacuum tubes (as a result of complementary assets from the tube era that were valuable in the transistor era) and status as an early entrant (most of which were vacuum tube producers) are most advantageous with respect to firm performance. Since firm background played such a large role in firm performance in the transistor era, clusters of transistor producers built up in regions that housed existing electronic clusters, further reinforcing the importance of these regions. Performance during the integrated circuit era, however, appears to have relied on previous experience in planar transistors, production at the technological frontier, and a firm's status as a spinoff firm. As spinoffs from Fairchild and other firms used the knowledge gained in previous employment to their advantage

and began exhibiting superior performance, Silicon Valley started to emerge as the leading integrated circuit cluster.

The emergence of the technological frontier and spinoff status as important attributes with regard to performance appears to be linked to the nature of the technology that was required to produce integrated circuits at the technological frontier (monolithic integrated circuits). As explained, the technology required for monolithic integrated circuits was much more complex and tacit than that required to produce even the most complex transistors, though similarities between the technologies existed. Through the analyses performed, it was apparent that a background in silicon transistor production was important to being able to produce products at the technological frontier upon entry, presumably because of the similarities between existing silicon transistor production technologies (planar) and monolithic integrated circuit production. Silicon transistor firms already had access to a good deal of the complex, tacit knowledge involved in the monolithic integrated circuit production processes, and they were able to use this access to produce at the technological frontier upon entry. Following entry, very few firms (less than 10 percent) were able to adopt production of products at the technological frontier, and those that did so were primarily transistor firms, again utilizing knowledge from the transistor production process. As a result, access to this knowledge and the ability to apply this knowledge to produce integrated circuit products was vital for firms that were looking to succeed in this era.

Using detailed production data and new spinoff/parent data, this dissertation shows that spinoff firms are more likely to produce a given product if their parent firm was actively producing it at the time the spinoff was founded. The large number of spinoffs within Silicon Valley, most of which in some way were tied to the pioneering firm Fairchild, allowed for this knowledge

dissemination, accounting for the higher proportion of firms within the region producing at the technological frontier and ultimately the leadership of the region with respect to firm performance during the integrated circuit era. Ultimately, spinoffs provided access to knowledge regarding the technological frontier in ways that knowledge spillovers within the region could not.

Results suggest that regional knowledge spillovers and other agglomeration effects may not be overly important to the development of clusters. No findings from either era suggest that any regional effects were present once the background and technological frontier status of producers were included in analyses examining either survival or market share attainment. This non-existence of regional effects may be particularly evident when the knowledge required for production is complex and tacit. In this case, knowledge is difficult to disseminate to others not directly involved in the production process, and thus individuals need to learn by doing, an opportunity afforded to industry employees that later decide to form their own firms. It should be noted, however, that spinoff firms do not seem overly important during the transistor era, indicating that the nature of the technology is important when considering how firms will gain access to knowledge and how the industry will evolve.

Given the role of the nature of technology in how cutting-edge knowledge is disseminated, the implications of the nature of the technology on the design of policies and programs for regional development are important and clear. If the technology is relatively open or able to be disseminated in ways that do not require direct experience, regional cluster development strategies focused on existing producers with related experience may be appropriate and effective. However, in the case that knowledge is complex and tacit, and complementary assets

of existing producers are not valuable, these strategies may need to be modified to rely more on mechanisms that expose potential entrepreneurs to direct experience with cutting-edge technologies. Any efforts to build regional clusters when knowledge is complex and tacit may benefit from facilitating employee mobility in the form of spinoff firms, especially the mobility of those individuals working at high quality firms who possess the most valuable product, market and management knowledge. While the elimination of non-compete agreements is one approach to increasing mobility, other approaches such as increased networking opportunities or seed grants to fund start-ups among individuals with previous industry experience may also yield positive results. Important to this approach, however, is that access to the appropriate leading edge knowledge is vital in the creation of successful firms.

The existence of knowledge within leading firm(s) is important because it helps to fuel the development of high quality spinoffs. As we saw in the integrated circuit era, it was mostly the existence of high quality spinoff firms that fueled the incredible growth and performance in Silicon Valley, rather than a blanket regional effect. The creation of spinoff firms may be possible, at a fundamental level, by the fact that non-compete employment contracts were not enforceable in California (Fallick, Fleischman, & Rebitzer, 2008), unlike many other states (Gilson, 1998). However, if the creation of spinoff firms was solely based on employee mobility, one would expect Los Angeles to have developed in a similar fashion. Rather, high-quality spinoffs were founded by individuals who had previously worked for a high-quality parent company. These high-quality companies existed in much larger numbers in Silicon Valley as compared to other regions, partially due to the internal turmoil at Fairchild which then

resulted in a larger number of spinoffs (and high-quality firms) there with respect to other regions.[48]

As governments design policies and programs to encourage regional economic development, it is clear that the nature of the technology involved in the industry needs to be considered. Cluster development cannot be seen as a one size fits all approach, and what works for one region or one industry may not be effective in another setting. Given the large amount of resources that are currently being invested in regional cluster development, this insight which was generated through the analyses of this dissertation is potentially very valuable.

[48] 3 in Boston, 1 in Los Angeles, 2 in New York City, 20 in Silicon Valley

References

Address by Mr. James M. Bridges, Director, Office of Electronics, Office of the Director of Defense Research & Engineering, Before the Ft. Monmouth Chapter of the Armed Forces Communications Electronics Association. (1963). Texas Instruments records, DeGolyer Library, Southern Methodist University.

Aitken, H. G. J. (1994). Allocating the Spectrum: The Origins of Radio Regulation. *Technology and Culture*, *35*(4), 686-716.

Alcacer, J., & Chung, W. (2007). Location Strategies and Knowledge Spillovers. *Management Science*, *53*(5), 760-776. INFORMS. doi:10.1287/mnsc.1060.0637

Almeida, P., & Kogut, B. (1999). Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science*, *45*(7), 905 - 917. INFORMS. Retrieved from http://www.jstor.org/stable/2661691

Anton, J. J., & Yao, D. A. (1995). Start-ups, Spin-offs, and Internal Projects. *Journal of Law, Economics and Organization*, *11*(2), 362-378.

Audretsch, D. B., & Feldman, M. P. (1996). R&D Spillovers and the Geography of Innovation and Production. *The American Economic Review*, *86*(3), 630 - 640. American Economic Association. Retrieved from http://www.jstor.org/stable/2118216

Baptista, R., & Swann. (1998). Do firms in clusters innovate more? *Research Policy*, *27*(5), 525-540. doi:10.1016/S0048-7333(98)00065-1

Basic Electrical And Electronics Engineering (PTU, Jalandhar). (2006). (p. 760). Firewall Media. Retrieved from http://books.google.com/books?id=UOvVUw3OiTQC&pgis=1

Bassett, R. K. (2002). *To the digital age: research labs, start-up companies, and the rise of MOS Technology* (p. 421). JHU Press.

Braun, E., & Macdonald, S. (1982). *Revolution in miniature: the history and impact of semiconductor electronics* (p. 247). Cambridge University Press.

Breschi, S., & Lissoni, F. (2006). Mobility of inventors and the geography of knowledge spillovers. New evidence on US data. *CESPRI Working Papers 184*.

Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography*, *9*(4), 439-468. doi:10.1093/jeg/lbp008

Bresnahan, T. F., & Gambardella, A. (2004). *Building high-tech clusters: Silicon Valley and beyond* (p. 369). Cambridge University Press.

Brock, D. C. (2006). Reflections on Moore's Law. *Understanding Moore's Law: four decades of innovation*.

Brower, W. C., Cragon, H. G., & Lathrop, J. W. (1965). *Final Report on Silicon Semiconductor Networks Manufacturing Methods, Report Number AFML-TR-64-386* (pp. 87-27, Bow 3 Folder 1). Dallas, TX. Texas Instruments records, DeGolyer Library, Southern Methodist University.

Buenstorf, G., & Klepper, S. (2009). Heritage and Agglomeration: The Akron Tyre Cluster Revisited. *The Economic Journal*, *119*(537), 705-733. doi:10.1111/j.1468-0297.2009.02216.x

Burgess, M. P. D. (2008). RCA Transistor History. Retrieved July 2, 2011, from http://sites.google.com/site/transistorhistory/Home/us-semiconductor-manufacturers/rca-history

Burgess, M. P. D. (2011a). Semiconductor Research and Development at General Electric. Retrieved July 2, 2011, a from http://sites.google.com/site/transistorhistory/Home/us-semiconductor-manufacturers/general-electric-history

Burgess, M. P. D. (2011b). Early Semiconductor History of Texas Instruments. Retrieved July 11, 2011, b from http://sites.google.com/site/transistorhistory/Home/us-semiconductor-manufacturers/ti

Carlton, D. W. (1983). The Location and Employment Choices of New Firms: An Econometric Model with Discrete and Continuous Endogenous Variables. *The Review of Economics and Statistics*, *65*(3), 440 - 449. The MIT Press. Retrieved from http://www.jstor.org/stable/1924189

Cassiman, M., & Ueda, B. (2006). Optimal Project Rejection and New Firm Start-ups. *Management Science*, *52*(2), 262-275.

Chen, J. (2004). *Study of Germanium MOSFETs With Ultrathin High-k Gate Dielectrics*. University of Texas at Austin.

Choi. (2007). The Boundaries of Industrial Research: Making Transistors at RCA, 1948–1960. *Technology and Culture*, *48*(4), 758.

Choi, H., & Mody, C. C. M. (2009). The Long History of Molecular Electronics: Microelectronics Origins of Nanotechnology. *Social Studies of Science*, *39*(1), 11-50. doi:10.1177/0306312708097288

Christensen, C. M. (1997). *The innovator's dilemma: the revolutionary book that will change the way you do business*.

Chung, W., & Alcacer, J. (2002). Knowledge seeking and location choice of Foreign Direct Investment in the United States. *Management Science*, 1534-1554.

City of San Jose. (n.d.). Retrieved February 20, 2012, from http://www.sanjoseca.gov/about.asp

Classification of Integrated Circuits. (n.d.). Retrieved January 6, 2012, from http://www.circuitstoday.com/classification-of-integrated-circuits

Climate of Moscow. (n.d.). Retrieved May 2, 2012, from http://en.wikipedia.org/wiki/Climate_of_Moscow

Combes, P., & Duranton, G. (2006). Labour pooling, labour poaching, and spatial clustering. *Regional Science and Urban Economics*, *36*(1), 1-28. doi:10.1016/j.regsciurbeco.2005.06.003

Computer History Museum - The Silicon Engine | 1958 - Silicon Mesa Transistors Enter Commercial Production. (n.d.). Retrieved February 25, 2012, from http://www.computerhistory.org/semiconductor/timeline/1958-Mesa.html

Computer History Museum - The Silicon Engine | 1960 - First Planar Integrated Circuit is Fabricated. (n.d.). Retrieved February 25, 2012, from http://www.computerhistory.org/semiconductor/timeline/1960-FirstIC.html

Computer History Museum - The Silicon Engine | 1962 - Aerospace systems are first the applications for ICs in computers. (n.d.). Retrieved February 26, 2012, from http://www.computerhistory.org/semiconductor/timeline/1962-Apollo.html

Computer History Museum - The Silicon Engine | 1965 - Package is the First to Accommodate System Design Considerations. (n.d.). Retrieved January 8, 2012, from http://www.computerhistory.org/semiconductor/timeline/1965-Package.html

Cooper, A. C., & Schendel, D. (1976). "Strategic responses to technological threats." *Business Horizons*, *19*(1), 61.

Cooper, D. (2001). Innovation and reciprocal externalities: information transmission via job mobility. *Journal of Economic Behavior & Organization*, *45*(4), 403-425. doi:10.1016/S0167-2681(01)00154-8

Dahl, M. S., & Sorenson, O. (2009). The embedded entrepreneur. *European Management Review*, *6*(3), 172-181. doi:10.1057/emr.2009.14

Doremus, J. A. (1952, April). Point-Contact and Junction Transistors. *Radio & Television News*.

Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics, Vol. 4: Cities and Geography*. Amsterdam, Netherlands: Elsevier.

Electronic Industries Association Marketing Services Department. (1980). *Electronic Market Data Book*. Washington, DC.

European Cluster Alliance. (n.d.). Retrieved February 5, 2012, from http://www.proinno-europe.eu/eca

Evans, R. A. (1998). Electronics reliability: a personal view. *IEEE Transactions on Reliability*, *47*(3), SP329-SP332. doi:10.1109/24.740548

Fallick, B., Fleischman, C. A., & Rebitzer, J. B. (2006). Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster. *Review of Economics and Statistics*, *88*(3), 472-481. The MIT Press. doi:10.1162/rest.88.3.472

Figueiredo, O., Guimaraes, P., & Woodward, D. (2002). Home-field advantage: location decisions of Portuguese entrepreneurs. *Journal of Urban Economics*, *52*(2), 341-361. doi:10.1016/S0094-1190(02)00006-2

Fleming, L., & Frenken, K. (2007). The evolution of inventor networks in the Silicon Valley and Boston regions. *Advances in Complex Systems*, *1*, 53-71.

Fleming, L., & Marx, M. (2006). Managing Creativity in Small Worlds. *California Management Review*, *48*(4), 6-27.

Fosfuri, A., & Ronde, T. (2004). High-tech clusters, technology spillovers, and trade secret laws. *International Journal of Industrial Organization*, *22*(1), 45-65. doi:10.1016/S0167-7187(03)00123-1

Franco, A. M., & Filson, D. (2006). Spin-outs: knowledge diffusion through employee mobility. *The RAND Journal of Economics*, *37*(4), 841-860. doi:10.1111/j.1756-2171.2006.tb00060.x

Fujita, M., Krugman, P., & Venables, A. (2001). *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.

Gilson, R. J. (1998). The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete. *SSRN Electronic Journal*. SSRN. doi:10.2139/ssrn.124508

Goldstein, A. (1991). Gordon K. Teal: An Interview Conducted by Andrew Goldstein. *IEEE History Center*. Retrieved July 2, 2011, from http://www.ieeeghn.org/wiki/index.php/Oral-History:Gordon_K._Teal

Haggerty, P. (1963). Memo (September 10, 1963) to John R. Moore (Autonetics). Texas Instruments records, DeGolyer Library, Southern Methodist University.

Hall, R., & Dunlap, W. (1950). P-N Junctions Prepared by Impurity Diffusion. *Physical Review*, *80*(3), 467-468. American Physical Society. doi:10.1103/PhysRev.80.467.2

Hamilton, M. (2008). *Exploring Intervention to the Regional Institutions for Innovation: Technology-based Economic Development*. Carnegie Mellon University.

Harrigan, K. R., & Porter, M. (1979). *The Receiving Tube Industry in 1966*. Boston, MA: HBS Case Services, Harvard Business School.

Harrison, A. P. J. (1979). Single-Control Tuning: An Analysis of an Innovation. *Technology and Culture*, *20*(2), 296-321.

Head, K., Ries, J., & Swenson, D. (1995). Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States. *Journal of International Economics*, *38*(3-4), 223-247. doi:10.1016/0022-1996(94)01351-R

Helsley, R., & Strange, W. C. (1990). Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics*, *20*(2), 189-212. doi:10.1016/0166-0462(90)90004-M

Henderson, J. (1998). Evidence on Scale Economies and Agglomeration. *Brown University Working Paper*.

Henderson, J. (2003). Marshall's scale economies. *Journal of Urban Economics*, *53*(1), 1-28. doi:10.1016/S0094-1190(02)00505-3

Henderson, J. V. (1974). The Sizes and Types of Cities. *The American Economic Review*, *64*(4), 640 - 656. American Economic Association. Retrieved from http://www.jstor.org/stable/1813316

Henderson, R., & Clark, K. B. (1990). Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms. *Administrative Science Quarterly*, *35*(1), 9-30. Retrieved from http://www.jstor.org/stable/10.2307/2393549

Herold, E. W. (1983). Focus on a Career Engineer - Excerpts from Bygone Days: An Autobiography. Retrieved July 2, 2011, from http://www.davidsarnoff.org/ewh.html

Heyer, M., & Pinsky, A. (1975). Dr. Charles W. Mueller, Electrical Engineer, an oral history. *IEEE History Center*. Retrieved from http://www.ieeeghn.org/wiki/index.php/Oral-History:Charles_W._Mueller#Interview

Hoefler, D. (1971a, January 18). Silicon Valley U.S.A. *Electronic News*, pp. 1, 4-5.

Hoefler, D. (1971b, January 11). Silicon Valley U.S.A. *Electronic News*, pp. 1, 4-5.

Hoefler, D. (1971c, January 25). Silicon Valley U.S.A. *Electronic News*, pp. 4-5.

Holbrook, D. (1995). Government Support of the Semiconductor Industry: Diverse Approaches and Information Flows. *Business and Economic History*, *24*(2), 133-165.

Holbrook, D. (1999). *Technical diversity and technological change in the American semiconductor industry, 1952-65*. Carnegie Mellon University.

Hybrid integrated circuit - Wikipedia, the free encyclopedia. (n.d.). Retrieved January 6, 2012, from http://en.wikipedia.org/wiki/Hybrid_integrated_circuit

Kabaservice, T. P. (1978). *Applied Microelectronics*. St. Paul, MN: West Publishing Co.

Kaplan, D. (2000). *The Silicon Boys: And Their Valley of Dreams*. New York, NY: Harper Perennial.

Kilby, J. S. (1976). Invention of the integrated circuit. *IEEE Transactions on Electron Devices*, *23*(7), 648-654. doi:10.1109/T-ED.1976.18467

Klepper, S. (2007a). Disagreements, Spinoffs and the Evolution of Detroit at the Capital of the U.S. Automobile Industry. *Management Science, 53*(4), 616-631.

Klepper, S. (2007b). The Geography of Organizational Knowledge.

Klepper, S. (2009a). Silicon Valley — A Chip off the Old Detroit Bloc. In D. B. Audretsch & R. Strom (Eds.), *Entrepreneurship, Growth and Public Policy*. Cambridge, England: Cambridge University Press.

Klepper, S. (2009b). Spinoffs: A review and synthesis. *European Management Review, 6*(3), 159-171. doi:10.1057/emr.2009.18

Klepper, S. (2010). The origin and growth of industry clusters: The making of Silicon Valley and Detroit. *Journal of Urban Economics, 67*(1), 15-32. doi:10.1016/j.jue.2009.09.004

Klepper, S., & Simons, K. L. (2000). Dominance by birthright: entry of prior radio producers and competitive ramifications in the U.S. television receiver industry. *Strategic Management*

Journal, *21*(10-11), 997-1016. doi:10.1002/1097-0266(200010/11)21:10/11<997::AID-SMJ134>3.0.CO;2-O

Klepper, S., & Sleeper, S. (2005). Entry by Spinoffs. *Management Science*, *55*, 1291-1306.

Klepper, S., & Thompson, P. (2010). Disagreements and intra-industry spinoffs. *International Journal of Industrial Organization*, *28*(5), 526-538. doi:10.1016/j.ijindorg.2010.01.002

Kowalski, J. (2011a). Interview with Ross Macdonald.

Kowalski, J. (2011b). Interview with Harvey Cragon.

Lane, S. (1989). *Entry and industry evolution in the ATM manufacturers' market*. Stanford University.

Lathrop, J. W., Lee, R. E., & Phipps, C. (1960, May 13). Semiconductor Networks for Microelectronics. *Electronics*.

Lécuyer, C. (1999). Silicon for industry: Component design, mass production, and the move to commercial markets at Fairchild semiconductor, 1960-1967. *History and Technology*, *16*(2), 179-216.

Lécuyer, C. (2000). Fairchild Semiconductor and Its Influence. In C.-M. Lee, W. F. Miller, M. G. Hancock, & H. S. Rowen (Eds.), *The Silicon Valley Edge: A Habitat for Innovation and Entrepreneurship* (1st ed., pp. 158-183). Stanford, CA: Stanford University Press.

Lécuyer, C. (2006). *Making Silicon Valley: innovation and the growth of high tech, 1930-1970* (p. 393). Cambridge, MA: MIT Press.

Lécuyer, C., & Brock, D. C. (2010). *Makers of the Microchip: A Documentary History of Fairchild Semiconductor*. Cambridge, MA: MIT Press.

Leslie, S. W., & Kargon, R. H. (1996). Selling Silicon Valley: Frederick Terman's Model for Regional Advantage. *The Business History Review*, *70*(4), 435-472.

Lieberman, M. B., & Montgomery, D. B. (1988). First-mover advantages. *Strategic Management Journal*, *9*(S1), 41-58. doi:10.1002/smj.4250090706

Macdonald, J. R. (2009). Employment for James Ross Macdonald. Retrieved July 16, 2011, from http://www.jrossmacdonald.com/About-JRM/jobs.html

Malone, M. (1985). *The Big Score: The billion Dollar Story of Silicon Valley*. Garden City: Doubleday.

Manufacture of Junction Transistors. (n.d.). Retrieved January 6, 2012, from http://www.thevalvepage.com/trans/manufac/manufac1.htm

Marshall, A. (1890). *Principles of Economics: an introductory volume*. London, UK: Macmillan & Co.

Marx, M., Strumsky, D., & Fleming, L. (2009). Mobility, Skills, and the Michigan Non-Compete Experiment. *Management Science, 55*(6), 875-889. doi:10.1287/mnsc.1080.0985

Matthews, G. (2003). *Silicon valley, women, and the California dream: gender, class, and opportunity in the twentieth century* (p. 313). Stanford University Press. Retrieved from http://books.google.com/books?id=qbBfBukuMMwC&pgis=1

Merryman, S. L. (1988). Application For An Historical Marker Commemorating The Demonstration Of the First Working Integrated Circuit. Dallas, TX: Texas Instruments records, DeGolyer Library, Southern Methodist University.

Michelacci, C., & Silva, O. (2007). Why So Many Local Entrepreneurs? The MIT Press. Retrieved from http://www.mitpressjournals.org/doi/abs/10.1162/rest.89.4.615

Millis, E. (2000). *TI, the Transistor, and Me* (1st ed.). Dallas, TX: Texas Legal Copies.

Mindell, D. (2008). *Digital Apollo*. Cambridge, MA: MIT Press.

Moore. (1998). The Role of Fairchild in Silicon Technology in the Early Days of "Silicon Valley." *Proceedings of the IEEE, 86*(1), 53-62.

Moore, G., & Davis, K. (2004). Learning the Silicon Valley Way. In T. F. Bresnahan & A. Gambardella (Eds.), *Building high-tech clusters: Silicon Valley and beyond*. Cambridge, England: Cambridge University Press.

Morgan, J. (1967). *Electronics in the West: The First Fifty Years*. Palo Alto, CA: National Press Books.

Morris, P. R. (1990). *A History of the World Semiconductor Industry*. London, UK: Peter Peregrinus Ltd.

Museum, C. H. (n.d.). 1955 - Photolithography Techniques Are Used to Make Silicon Devices. Retrieved July 7, 2011, from http://www.computerhistory.org/semiconductor/timeline/1955-Photolithography.html

Nelson, R. R. (1959). The Simple Economics of Basic Scientific Research. *Journal of Political Economy, 67*(3), 297-306.

Nelson, R. R. (1962). The Link Between Science and Invtion: The Case of the Transistor. In U.-N. Bureau (Ed.), *The Rate and Direction of Inventive Activity: Economic and Social Factors* (pp. 549-584).

Pakes, A., & Nitzan, S. (1983). Optimum Contracts for Research Personnel, Research Employment, and the Establishment of "Rival" Enterprises. *Journal of Labor Economics, 1*(4), 345 - 365. The University of Chicago Press on behalf of the Society of Labor Economists and the National Opinion Research Center. Retrieved from http://www.jstor.org/stable/2534859

Parwada, J. T. (2008). The Genesis of Home Bias? The Location and Portfolio Choices of Investment Company Start-Ups. *Journal of Financial and Quantitative Analysis, 43*(01), 245-266. Retrieved from http://journals.cambridge.org/abstract_S0022109000002817

Pearson, G. L., & Brattain, W. H. (1955). History of Semiconductor Research. *Proceedings of the IEEE, 43*(12), 1794-1806.

Phipps, C., & Laws, D. (2011). The Early History of Integrated Circuits at Texas Instruments: A Personal View. *IEEE Annals of the History of Computing*, (99), 1-1. doi:10.1109/MAHC.2011.84

Physical Fabrication of Transistors. (n.d.). Retrieved February 26, 2012, from http://www.cjseymour.plus.com/elec/basicfab/fab.htm

Porter, M. (2001). Regions and the New Economics of Competition. In A. J. Scott (Ed.), *Global City-Regions* (pp. 139-157). New York: Oxford University Press.

Reid, T. R. (2001). *The Chip: How Two Americans Invented the Microchip and Launched a Revolution* (2nd ed.). Random House Trade Paperbacks.

Riordan, M. (2004). The Lost History of the Transistor. *IEEE Spectrum*. Retrieved July 16, 2011, from http://spectrum.ieee.org/biomedical/devices/the-lost-history-of-the-transistor

Riordan, M. (2007). From Bell Labs to Silicon Valley: A Saga of Semicondcutor Technology Transfer, 1955-61. *The Electrochemical Society Interface*, (Fall), 36-41.

Riordan, M., & Hoddeson, L. (1997). *Crystal Fire*. New York: W.W. Norton & Company.

Saddler, I. (1990). The Human Side of Early Electronics and Semiconductors. *SMEC Vintage Electronics, 2*.

Saxenian, A. (1994). *Regional advantage: culture and competition in Silicon Valley and Route 128* (p. 226). Harvard University Press.

Scott-Morton, F. (1999). Entry Decisions in the Generic Pharmaceutical Industry. *RAND Journal of Economics, 30*, 421-440.

Seidenberg, P. (1997). From Germanium to Silicon: A History of Change in the Technology of Semiconductor. In A. Goldstein & W. Aspray (Eds.), *Facets: New Perspectives on the History of Semiconductors* (pp. 35-74). New Brunswick, NJ: IEEE Center for the History of Electrical Engineering.

Shockley, W. (1950). *Electrons and Holes in Semiconductors* (New York.). D. Van Nostrand Co. Inc.

Shurkin, J. N. (2006). *Broken Genius - The rise and fall of William Shockley, creator of the electronic age*. New York, NY: Macmillan.

Silicon Valley Genealogy Chart. (1995). Semiconductor Equipment and Materials International.

Skokan, K. (2011). Financing the Cluster Development in Moravia Silesia by EU Funds. *8th International scientific conference Financial management of firms and finiancial institutions*. Ostrava.

Skolkovo Innovation Center: Executive Summary. (2010). Moscow. Retrieved from http://www.slideshare.net/iponomarev/100930-booz-mm-skolkovo-executive-summary

Skolkovo Participant Benefits. (n.d.). Retrieved February 5, 2012, from http://www.sk.ru/GetInvolved/Innovator/ParticipantsBenefits.aspx

Stuhlbarg, S. M. (1961). How Industry Sees Microminiaturization. Texas Instruments records, DeGolyer Library, Southern Methodist University. Retrieved from 86-10 Semi Conductor Networks 1961

Sturgeon, T. J. (2000). How Silicon Valley Came to Be. *Understanding Silicon Valley: The Anatomy of an Entrepreneurial Region* (pp. 15-47). Stanford, CA: Stanford University Press.

Tanenbaum, M. (2008). First Hand: Beginning of the Silicon Age. *IEEE Global History Network*. Retrieved July 12, 2011, from http://ghn.ieee.org/wiki/index.php/First-Hand:Beginning_of_the_Silicon_Age

Texas Instruments - 1954 first commercial silicon transistor. (n.d.). Retrieved February 20, 2012, from http://www.ti.com/corp/docs/company/history/timeline/semicon/1950/docs/54commercial.htm

Texas Instruments - Transistor History. (n.d.). Retrieved February 26, 2012, from http://sites.google.com/site/transistorhistory/Home/us-semiconductor-manufacturers/ti

Texas Instruments, C. I. D. (1964). Likely Editorial Questions upon Receipt of June 23 News Release. Texas Instruments records, DeGolyer Library, Southern Methodist University. Retrieved from 85-24, Box 4 Folder 3 PR, 86-10 Semiconductor Networks 1966

The Chip that Jack Built. (n.d.). Retrieved January 6, 2012, from http://www.ti.com/corp/docs/kilbyctr/jackbuilt.shtml

The Hybrid Microcircuit Micromodule and Solid Logic Technology. (n.d.). Retrieved January 6, 2012, from http://www.chipsetc.com/the-hybrid-microcircuit.html

Tilton, J. (1971). *International Diffusion of Technology: The Case of Semiconductors*. Washington, DC: Brookings Institution.

Tripsas, M. (1997). Unraveling the process of creative destruction: Complementary assets and incumbent survival in the typesetter industry. *Strategic Management Journal*, *18*(Sp. Iss. SI), 119-142. JOHN WILEY & SONS LTD. Retrieved from http://apps.isiknowledge.com/full_record.do?product=UA&search_mode=GeneralSearch&qid=4&SID=2F6iGgMnIE8Hkf@Kgjn&page=1&doc=1&colname=WOS

Tushman, L. M., & Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative Science Quarterly*, *31*(3), 439-465. Retrieved from http://www.jstor.org/stable/10.2307/2392832

Unknown. (n.d.-a). *Innovation Insights: The IC Story*. Retrieved from http://www.youtube.com/watch?v=DmizzbsdXfc

Unknown. (n.d.-b). S/C Network Manufacturing Yields. Texas Instruments records, DeGolyer Library, Southern Methodist University.

Unknown. (n.d.-c). Semiconductor Network Processes. Texas Instruments records, DeGolyer Library, Southern Methodist University.

Vintage Semiconductors Ltd Transistors. (n.d.). Retrieved February 26, 2012, from http://www.wylie.org.uk/technology/semics/Semics/Semics.htm

Ward, J. (2001). Early Transistor History at Texas Instruments: Elmer Wolff Jr. Retrieved July 13, 2011, from http://semiconductormuseum.com/Transistors/TexasInstruments/OralHistories/Wolff/Wolff_Index.htm

What is Skolkovo? (n.d.). Retrieved February 8, 2012, from http://www.sk.ru/en/Model.aspx

Why Dmitry Medvedev wants a Russian Silicon Valley. (n.d.). Retrieved February 10, 2012, from http://www.smartplanet.com/blog/thinking-tech/why-dmitry-medvedev-wants-a-russian-silicon-valley/4531

Willis Adcock - Wikipedia. (2011). Retrieved July 16, 2011, from http://en.wikipedia.org/wiki/Willis_Adcock

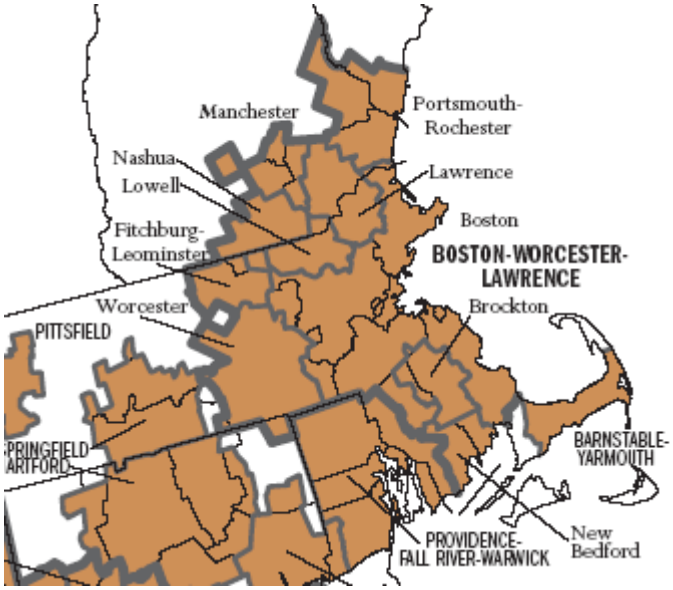
Wolff, M. (1985). Roger Webster: An Interview Conducted by Michael Wolff. New Brunswick, NJ: IEEE History Center.

Zygmont, J. (2003). *Microchip: An idea, its genesis, and the revolution it created*. Cambridge, MA: Perseus Publishing.

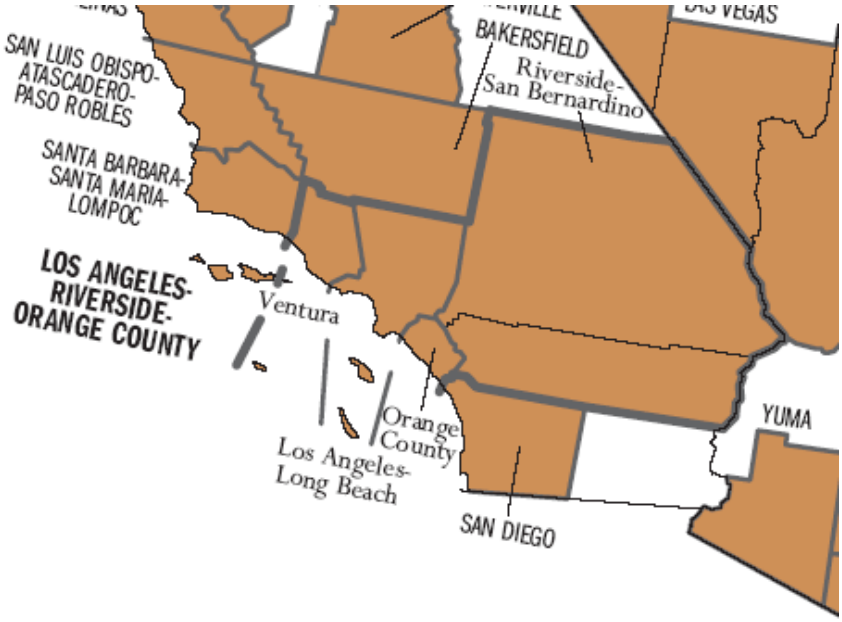
von Hippel, E. (1987). Cooperation between rivals: Informal know-how trading. *Research Policy*, *16*(6), 291-302. doi:10.1016/0048-7333(87)90015-1

Appendix A: CMSA Definition Maps

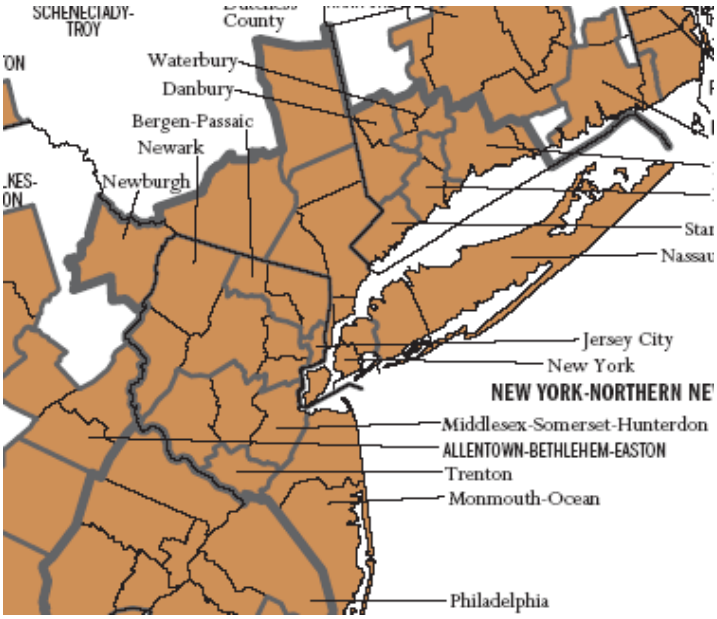
Boston Map



Los Angeles Map



New York City Map



Silicon Valley (San Francisco) Map

Source: US Census Bureau (http://www.census.gov/geo/www/mapGallery/ma_1999.pdf)

Appendix B: Product Categories

Table 27: Broad Integrated Circuit Product Categories

Broad Product Category	First Production Year	Last Production Year	Peak Production Year	Firms Producing in Peak Year
Film	1965	1987	1972	116
Hybrid	1965	1987	1969	145
Monolithic Custom	1970	1987	1987	67
Monolithic Linear	1970	1987	1987	97
Monolithic Logic	1970	1987	1975	62
Monolithic Memory	1970	1987	1987	45
Monolithic Microprocessor	1976	1987	1987	52
Monolithic Other	1965	1987	1987	84

Table 28: Narrow Integrated Circuit Product Categories

Product Code	Product Description	First Production Year	Last Production Year	Peak Production Year	Firms Producing in Peak Year
FTHICK	Thick Film	1969	1987	1987	87
FTHIN	Thin Film	1965	1987	1972	100
HCDIG	Hybrid Custom Digital Logic	1965	1987	1969	82
HCEC	Hybrid Consumer Electronics Circuits	1973	1987	1987	25
HCL	Hybrid Linear Circuits	1965	1969	1969	74
HCONV	Hybrid Converters A/D & D/A	1977	1987	1987	33
HDD	Hybrid Digital Drivers	1973	1987	1975	34
HIC	Hybrid Instrumentation & Control	1973	1987	1987	34
HLSI	Hybrid Large Scale Integration	1968	1972	1972	52
HMCHIP	Hybrid Multiple Chip	1965	1972	1969	91
HMUX	Hybrid Multiplexers	1977	1987	1987	19
HPS	Hybrid Packaging Systems	1968	1969	1969	38
HRAD	Hybrid Radio	1973	1987	1974	32

	Communications				
HSTG	Hybrid Storage & Memory	1973	1987	1975	19
HSW	Hybrid Switches	1973	1987	1987	30
HVR	Hybrid Voltage or Current Regulators	1974	1987	1984	34
LCMOS	Monolithic Logic CMOS	1973	1987	1987	46
LDIG	Monolithic Digital Logic	1970	1976	1972	44
LDTL	Monolithic Logic DTL	1973	1987	1981	24
LECL	Monolithic Logic ECL	1973	1987	1982	18
LMOS	Monolithic Logic MOS	1973	1987	1975	41
LRTL	Monolithic Logic RTL	1973	1987	1975	12
LTTL	Monolithic Logic TTL	1973	1987	1974	34
MAUD	Monolithic Audio Circuits	1977	1987	1987	57
MBPL	Monolithic Bipolar	1970	1987	1975	40
MCGCC	Monolithic Character Generators & Code Converters	1973	1982	1975	25
MCL	Monolithic Custom Linear	1970	1987	1987	52
MCONV	Monolithic A/D Converters	1976	1987	1987	43
MCW	Monolithic Shift Registers	1970	1987	1972	38
MDI	Monolithic Dielectrically Isolated	1973	1975	1975	27
MDIG	Monolithic Digital	1970	1987	1972	28
MGC	Monolithic Games Circuits	1977	1982	1977	34
MLA	Monolithic Linear Amplifiers	1970	1987	1975	51
MLD	Monolithic Linear Line Drivers	1976	1987	1987	25
MLS	Monolithic Linear Subsystems	1973	1976	1975	28
MMAM	Monolithic Multiplexers and Matrixes	1970	1975	1972	42
MMOD	Monolithic Modems	1973	1987	1984	23
MMOS	Monolithic MOS	1965	1987	1987	58
MMP	Monolithic Microprocessors	1976	1987	1987	52
MMW	Monolithic Microwave	1986	1987	1987	37
MOPTO	Monolithic Optoelectronic	1973	1987	1976	19
MPROM	Monolithic Memory Programmable	1973	1987	1987	35
MRAM	Monolithic Memory Random Access	1976	1982	1982	30
MROM	Monolithic Memory Read	1976	1987	1982	34

	Only				
MRTV	Monolithic Radio TV & Communications	1973	1987	1982	40
MSBC	Monolithic Side Brazed Ceramic	1980	1987	1985	13
MSTG	Monolithic Memory Storage	1970	1987	1972	37
MTIME	Monolithic Timers	1976	1987	1982	26
MVR	Monolithic Voltage Regulators	1970	1987	1987	39
MVRS	Monolithic Voice Recognition & Synthesis	1983	1987	1987	10