Paper to be presented at
the DRUID16 20th Anniversary Conference
Copenhagen, June 13–15, 2016

Scaling Moore’s Wall: A Public-Private Partnership in Search of a Technological Revolution

Hassan Khan
Carnegie Mellon University
Engineering and Public Policy
hassnkhan@cmu.edu

David Hounshell
Carnegie Mellon University
Social and Decision Sciences
hounshell@cmu.edu

Erica Fuchs
Carnegie Mellon University
Engineering and Public Policy
erhf23@gmail.com

Abstract
The decline of corporate research and vertical disintegration of supply chains in many industries has led to an innovation ecosystem increasingly reliant on linkages between institutions, presenting new challenges for long-term technology development. Pre-commercial public-private research consortia offer one policy response, and yet the majority of past research has focused on public-private consortia created for short-term (1-3-years out) technology development and technology catch-up. Based on unprecedented access to archives of the Semiconductor Research Corporation (SRC), publicly available data, 50 semi-structured interviews, and participant observation, we examine how the Nanoelectronics Research Initiative (NRI) emerged in response to arguably the most significant presumptive anomaly of our time: the end of Moore’s Law. NRI aimed to bridge the semiconductor industry’s past 40 years of unprecedented technology development—captured by Moore’s Law—with a radically new (and, as of this writing, not-yet-discovered) technology that will maintain this development indefinitely. Building on a long history of collaborative university-industry research programs managed by the Semiconductor Research Corporation (SRC), we suggest the NRI played a coordinating role within the scientific community. Specifically, we show how NRI incorporated industry expertise in manufacturing and design to inform and shape academic research aimed at inventing a successor to CMOS technology. We
conclude by questioning the extent to which the effort was appropriately suited to the nature and importance of the end-of-Moore’s Law challenge and the extent to which lessons from NRI may be generalized to a broader set of industrial contexts requiring coordination to overcome major technological discontinuities.
Scaling Moore’s Wall: A Public-Private Partnership in Search of a Technological Revolution

1. Introduction

The structure of industrial R&D has undergone considerable change in recent decades with a shift away from centralized R&D operations toward an increasing reliance on external sources of innovation (Chesbrough, 2003; Mowery, 2009). However, these shifts may pose significant challenges for firms facing technology discontinuities in vertically disintegrated industries, especially when the underlying science to navigate through that discontinuity is unknown. One potential policy tool for coordinating long-term technology development needs across firms, government, and academia, especially during times of technological discontinuities, is public-private partnerships. We look specifically at an industry facing a technological discontinuity where the solution and supporting science are almost wholly unknown.

Our study shows how the NRI emerged in response to arguably the most significant presumptive anomaly (Constant, 1980) of our time: the rapidly approaching end of Moore’s Law. Our research suggests that although NRI’s funding was small in comparison to larger federal R&D programs, the NRI played a coordinating role within the scientific community. Specifically, we show how, building on a long history of managing industry-university research programs within the Semiconductor Research Corporation, NRI incorporated industry expertise in manufacturing and design to inform and shape academic research, all with the explicit objective of finding an equally fecund alternative switch to the CMOS FET. However, given that the NRI program was ongoing as of the terminal date of our study, we make no normative judgment about NRI’s success or failure in meeting its objectives. Finally, we discuss the extent to which lessons from NRI may be generalized to a broader set of industrial contexts requiring coordination in long-term technology development through technical discontinuities.

2. Background:

2.1 Technological Uncertainty and Scientific Communities

The changing structure of corporate R&D has been concomitant with an increasing reliance by firms on external sources for R&D (Mowery and Rosenberg, 1991; Chesbrough, 2003; Mowery, 2009). Scholars have noted an increase in inter-firm collaboration (Hagedoorn, 2002) and inter-institutional collaborations (NAS, 2007), including a dramatic increase in industry-academic collaboration (Powell & Grodal, 2005) coupled with the increasing role of publicly funded science in generating technology patents (Narin et al., 1997). This increased reliance by firms on external sources of R&D presents new challenges for absorbing and influencing externally occurring innovation, particularly in periods of high scientific and technological uncertainty.
A host of work has sought to understand how companies can search for and absorb new technological information from the outside. A central tenet of this work has been the need for firms to develop the internal capabilities necessary to leverage external knowledge networks (Cohen and Levinthal, 1990; Powell et al. 1996), which can be achieved through a variety of mechanisms, including funding basic research within the firm (Rosenberg, 1990), establishing externally facing teams (Ancona and Bresman, 2007), engaging in strategic hiring (Singh and Agarwal, 2010; Rosenkopf and Almeida, 2003), and building alliances (Mowery et al., 1996). Recent work explores opportunities for external agents not only to identify and exploit external knowledge but also to influence the direction of scientific activities from which they may benefit. This research has shown that managers can “prune and pair” social networks within their own organization to achieve organizational goals (Davis, 2015) and act as “cocktail hostesses” in carefully selecting and bringing parties together to generate new ideas (Lester and Piore, 2004). Building on this work, research has shown the potential for a similarly (and arguably far more) active role for the state in orchestrating technology development. Recent studies have attributed DARPA’s success in meeting military needs to the processes employed by its program managers in understanding emerging themes in the research community, matching these emerging themes to military needs, and connecting disconnected communities, including cultivating “novel collaborations” between star researchers (Fuchs, 2010; Colatat, 2013).

Influencing the direction of scientific activities, both within key groups of researchers as well as among the broader scientific and technological community may be particularly important for industries facing long-term technology challenges. There is a significant body of scholarship on the relationship between scientific communities and scientific directions (Kuhn, 1962; de Solla Price, 1963; Crane, 1972), with a related literature on the role of communities in shaping technology directions (Constant, 1980; Bijker et al., 1987; Dosi, 1982; Schaller, 2004). Knowledge is also known to flow horizontally across firm boundaries (Allen, 1984; von Hippel 1989), particularly within groups with shared past experiences (Breznitz, 2005) and within communities of practice (Brown and Duguid, 2001; Almeida and Kogut, 1999).

Divergence in belief among researchers (Garud and Rappa, 1994) or widespread belief about future technology limits (Constant, 1980; Henderson, 1995) can foment technological uncertainty. “Presumptive anomaly” is a concept conceived by Constant (1980), analogous to the “anomalies” employed by Kuhn (1962) to show what precedes scientific revolution, to instead analyze the structure of “engineering revolution” or “technological revolution.” In Constant’s model, paradigm shift, or technological revolution, occurs not through the accretion of anomalies but from the projection of “presumptive anomalies”—or pushing the limits of a technology in the mind’s eye to the point of failure of the existing technological paradigm, system, or artifact (Constant 1980.) Constant cited the turbojet
revolution as the result of a set of presumptive anomalies involving aircraft flying at such high altitudes, in such thin atmospheres, and at such high speeds that existing piston engine-powered aircraft would totally fail (Constant, 1980.)

During times of technological uncertainty, conferences (Garud, 2008), communities of experts (Haas, 1989; Rosenkopf and Tushman, 1998), platform leaders (Gawer and Cusumano, 2002); and embedded network agents (Fuchs, 2010; Davis 2015) can play important roles in technology direction-setting. Indeed, research has suggested that the greater the technological uncertainty, the greater the number of alliances a company will form (Eisenhardt and Schoonhoven, 1996; Sarkar, Echambadi, and Harrison, 2001; Rosenkopf and Tushman, 1998). Community-wide direction-setting can be particularly important for platform technologies, which require coordination of extensive stakeholders across the industry, including upstream suppliers and downstream users (Gawer and Cusumano, 2002; Adner, 2006; Boudreau, 2010). As Constant writes, “often the same or very similar anomalous perceptions separately realized by several individuals or groups call forth different alternatives” (Constant, 1980 pg 18). Furthermore, “…the process of radical change in any one system… requires translation of its consequences into the interest frame of each of many relevant communities and the persuasion of each of them that the overall gains to be had … outweigh the costs” (Constant, 1980 pg 12). Market-coordination mechanisms alone may be incapable of properly addressing challenges posed by technological discontinuities. During long periods of incremental change, firms are often unwilling or unable to maintain the “systems level knowledge” necessary to carry out—or even adapt to—radical architectural shifts in technology (Chesbrough and Kusunoki, 2001). With the decline over the last three decades in corporate R&D, firms (and technological communities more broadly) may particularly lack the knowledge and resources necessary to overcome a technological grand challenge (NAE, 2015).

2.2 Public Private Partnerships and Technology Development

A host of organizational forms have been used in the past to address large-scale technology challenges (Corey, 1997; NRC, 2001; Powell and Grodal, 2005; CRS, 2009). In this paper we focus on one organizational form for coordinating long-term technology development needs across firms, government, and academia that has recently resurged in popularity: the public-private partnership. (See Appendix B.)

Past research on this organizational form has focused predominantly on public-private partnerships created to effect short-term technology development and technology catch-up (e.g. Spencer and Grindley, 1993; Mowery, 1998; Whitford, 2006). As a consequence, research on the processes for forming and carrying out public-private partnerships for addressing long-term technology development,
including during technological discontinuities and grand technology challenges, is needed. The gap in the literature on public-private partnerships for meeting technological grand challenges is particularly perplexing, given the acknowledged role for government in long-term technology development where uncertainties are high and the fundamental underlying science is unknown (Nelson, 2003).

The use of public-private partnerships to support industrial research and development efforts dates back to British “research associations” established late in World War I to try to bring British industry up to the research standards of Germany and the United States but were largely unsuccessful (Mowery and Rosenberg, 1991). Germany’s Kaiser Wilhelm Institutes demonstrated far more success in collaborative public-private research partnerships during the first four decades of the 20th-century (Beise and Stahl, 1999), and the Max Planck Institutes that succeeded them after World War II have also been deemed successful. In the U.S., public-private consortia focusing on pre-commercial research only began to form in the 1980s after changes to U.S. antitrust laws in response to the perceived successes of Japanese policymakers in the Ministry of International Trade and Industry (MITI) in guiding the restructuring of the Japanese computer and semiconductor industries with the use of publicly subsidized cooperative research programs (Sakakibara, 1993; Ouchi, 1984; Flamm, 1988; Sigurdson, 2004). Notably, however, many pre-commercial research consortia encountered difficulty in maintaining long-term horizons (Mowery and Teece, 1996). This was the case for the two most widely studied industry-government partnerships in the semiconductor industry: the VLSI program in Japan and SEMATECH in the U.S. Both consortia were geared towards helping national firms that had fallen behind international competition. The perceived result of both programs was improving the capabilities of semiconductor equipment supplier firms resulting in increased dependence by semiconductor manufacturers on national equipment suppliers (Sakakibara, 1993; Spencer and Grindley, 1993).

Traditionally, collaborative research has been framed in the academic literature as a mechanism for firms to internalize R&D externalities. Early theoretical literature focused on the level of knowledge spillovers and product competition as the main variables affecting the level and effectiveness of joint R&D expenditures (Katz, 1986; D’Aspremont and Jacquemin, 1988; Kamien et al, 1992). Empirical analyses of government-sponsored programs in Japan and the United States found evidence of a positive impact of consortium membership on participant firms' research productivity in the technology areas of collaboration (Sakakibara and Branstetter, 2003; Branstetter and Sakakibara, 2002). However, researchers have been unable to make strong causal claims about the impact of consortium membership on firm performance for large multinational firms (Link et al., 1996; Flamm and Wang, 2003). Additionally, as noted, research on the processes by which these public-private partnerships operate has been limited to programs focused on vertical (supplier and manufacturer) collaboration with short-term technology horizons (Whitford, 2006; Grindley et al., 1994).
More recently, research has argued that firms leverage collaborative research as a means for gaining strategic flexibility (Powell and Giannella, 2010), spreading costs of invention, accessing diverse sources of complementary knowledge (Mowery et al., 1996; Oxley and Wada, 2009), entering new business areas, and keeping track of technological frontiers (Sakakibara, 1997) – with specific reasons shaping their choice of collaborators (Belderbos et al., 2004). None of the above-cited scholarship focuses on overcoming technological grand challenges where the fundamental science is unknown. Powell and Giannella (2010) point out the theoretical importance of collective invention when there is uncertainty about the direction a technology will evolve and the kinds of applications that may emerge require broad search, but they do not offer in-depth empirical research on such cases. Thus, the research published to date does not offer insights into the processes by which public-private partnerships emerge and carry out such search, so as to orchestrate the scientific and technological frontier.

3. Methods

In the tradition of Rosenberg (1976), Henderson (1995), Bates et al. (1998), Levi (2002), Ingram et al. (2012) and Kahl et al. (2012) we develop an analytical narrative about the emergence and existence of the NRI. This analysis draws from three types of data: archival sources, semi-structured interviews, and participant observations (Jick, 1979). To date, we have conducted 50 interviews of member company researchers and executives, NRI researchers and center leaders, and SRC representatives to understand how those in the NRI envisage and assess its operations (see table 1 below). We have also conducted oral histories of individuals who played prominent roles in the formation and evolution of the industry’s key institutions including SRC, SEMATECH, and NTRS.

<table>
<thead>
<tr>
<th>Archival Data</th>
<th>Nanoelectronics Research Initiative (NRI)</th>
<th>Parallel Activities Outside NRI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inputs</strong></td>
<td>Firm financial data (Standard and Poor’s Compustat), primary source planning documents from NRI organizers</td>
<td>Government funding data: NSF funding archives, DARPA public funding archives (NSF, 2013; RDDS, 2013)</td>
</tr>
<tr>
<td><strong>Institutional</strong></td>
<td>NRI annual review presentation archives (2006-2012), SRC annual reports (1982-2013)</td>
<td>Key industry and government conferences (IEDM, DAC), Industry publications (EETimes, ACM)</td>
</tr>
<tr>
<td><strong>Individuals</strong></td>
<td>NRI executive and technical board representation, NRI conference attendees</td>
<td>Program directors at federal research funding agencies (ONR, DARPA, NSF, NIST)</td>
</tr>
<tr>
<td><strong>Outputs</strong></td>
<td>NRI research publications and patent filings</td>
<td>Patents and publications in NRI topic areas (U.S.PTO, 2013)</td>
</tr>
</tbody>
</table>

We conducted some 70 hours of participant observations of SRC and NRI conferences and annual research reviews to observe firsthand the organization’s deliberative processes. We attended the SRC’s
Techcon Conference in 2012 and the NRI’s Annual Reviews in 2012, 2013 and 2014. We also observed NRI Technical Program Group conference calls in the summer of 2013, during which NRI industry representatives discussed NRI administrative issues. Attending these conferences allowed us to witness firsthand how these programs operate and form our own assessments. SRC granted us considerable access to its archives, allowing us to analyze the historical development of SRC and providing historical context. Archival data utilized for this paper are listed in Table 2.

Table 2: Full list of interviews completed

<table>
<thead>
<tr>
<th>Name</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tom Theis</td>
<td>IBM/NRI</td>
</tr>
<tr>
<td>Jeff Welser</td>
<td>IBM/NRI</td>
</tr>
<tr>
<td>Paolo Gargini</td>
<td>Intel</td>
</tr>
<tr>
<td>George Scalise</td>
<td>SIA</td>
</tr>
<tr>
<td>Bob Doering</td>
<td>TI</td>
</tr>
<tr>
<td>George Bourianoff</td>
<td>Intel</td>
</tr>
<tr>
<td>Dimitri Nikonov</td>
<td>Intel</td>
</tr>
<tr>
<td>Steve Kramer</td>
<td>Micron</td>
</tr>
<tr>
<td>An Chen</td>
<td>GLOBALFOUNDRIES</td>
</tr>
<tr>
<td>Luigi Colombo</td>
<td>TI</td>
</tr>
<tr>
<td>Wilfried Haensch</td>
<td>IBM</td>
</tr>
<tr>
<td>Ajay Jacob</td>
<td>GLOBALFOUNDRIES</td>
</tr>
<tr>
<td>Zoran Krivokapic</td>
<td>GLOBALFOUNDRIES</td>
</tr>
<tr>
<td>Tony Low</td>
<td>IBM/Purdue</td>
</tr>
<tr>
<td>Tak Ning</td>
<td>IBM</td>
</tr>
</tbody>
</table>

Semiconductor Firms

SRC

<table>
<thead>
<tr>
<th>Name</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Larry Sumney</td>
<td>SRC</td>
</tr>
<tr>
<td>Ralph Cavin</td>
<td>SRC</td>
</tr>
<tr>
<td>Victor Zhirnov</td>
<td>SRC/ITRS</td>
</tr>
<tr>
<td>Steve Hillenius</td>
<td>SRC/GRC</td>
</tr>
<tr>
<td>Gil Vandentop</td>
<td>SRC/STARNet</td>
</tr>
<tr>
<td>Celia Merzbacher</td>
<td>SRC</td>
</tr>
</tbody>
</table>

Government

<table>
<thead>
<tr>
<th>Name</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mihail Roco</td>
<td>NSF/NNI</td>
</tr>
<tr>
<td>Lawrence Goldberg</td>
<td>NSF</td>
</tr>
<tr>
<td>Clifford Lau</td>
<td>DoD/ONR</td>
</tr>
<tr>
<td>Kerry Bernstein</td>
<td>DARPA</td>
</tr>
<tr>
<td>Larry Cooper</td>
<td>ONR</td>
</tr>
</tbody>
</table>
4.1 Emergence of a Public Private Partnership in Response to a Presumptive Anomaly

4.1.1 Anticipating Moore’s Wall

Over the course of some 50 years of integrated circuit manufacture, a series of presumptive anomalies were projected by members of the semiconductor industry with respect to the end of scaling of silicon transistor technology. Through 2004, the alternate technologies proposed as responses to the presumptive anomalies each time failed to overthrow the existing paradigm of scaling silicon devices, as technical advances extended the life of the prevailing technological system (Gilfillan, 1935; Rosenberg, 1976; Henderson, 1995).

As early as 1958 engineers at Westinghouse proposed a radical vision, “molecular electronics,” to officials in the Air Force as a way to leapfrog conventional silicon transistor technology (Holbrook, 1995; Choi and Mody, 2009). The program represented the first in a series of “presumptive anomaly”-inspired technologies proposed by the semiconductor industry that failed to supplant silicon. During the 1980s, researchers at the biggest industrial research labs in the industry actively used arguments about limits to
scaling to urge investment and research in alternative devices (e.g., Dennard, 1983; Heilmeier, 1984; Bate, 1986; Capasso et al., 1989). As firms scaled down the scope and size of their central research labs in the 1980s and 1990s their primary focus on incremental innovation for Si-CMOS advancement drove out long-range research on new-device structures. Long-range research on new device concepts, however, was kept alive by military research agency funding (Choi and Mody, 2009).

By the middle of the 1990s, the U.S. semiconductor industry had embraced the separation of design from manufacturing, and integrated device manufacturers had drastically cut the budgets of their central research labs and focused their research efforts toward support of their existing businesses (Macher et al., 1999). During the 1990s, the same institutions established by the industry to combat international competition and cooperate on advancing CMOS shifted the industry’s focus toward urgent long-term technology research needs and eventually toward the challenge of replacing CMOS.

Prior to 1994, the industry’s roadmapping exercises focused on 10-year extrapolations of the industry’s existing technology trajectory. In what eventually became known as Moore’s Law, industry pioneer Gordon Moore showed in 1965 that the number of components on a commercial integrated circuit had doubled roughly every 18 months and there were no reasons not to expect this trend to continue in the indefinite future (Moore, 1965). For CMOS transistors, the shrinking of feature sizes resulted in devices that were faster and more energy-efficient (Dennard et al., 1974).

The 1994 National Technology Roadmap for Semiconductors briefly identified three areas for long-term research: “nano-metrics metrology, nano-fabrication techniques, and new (post shrink) device structures” (NTRS, 1994 pg 6). The report outlined two parallel paths for “post-shrink” devices: (1) evolutionary concepts such as new geometries for the CMOS structure (e.g. dual gate and vertical MOSFETs) and (2) revolutionary “quantum device concepts” that had been pursued by the industry’s central research labs in the 1980s. The 1994 report set the scene for the coming evolution and potential revolution in microelectronics but highlighted the challenging long-term funding environment, noting that the decline of industrial research laboratories “[left] a major gap in the U.S. [research] infrastructure” (NTRS, 1994). An analysis commissioned by SIA and SRC argued that universities could address a considerable portion of the “research funding gap” (SRC, 1995).

The 1997 edition of the National Technology Roadmap for Semiconductors highlighted the challenge of adequately funding long-term research. The report’s primary foci were the challenging economics of continued scaling and impending challenges of integrating new materials throughout the CMOS architecture. In the decade following the 1997 report, the semiconductor industry devoted considerable engineering effort to overhauling the materials of the CMOS architecture in an effort to continue scaling. These new material solutions were engineered to mitigate complications from quantum effects at
increasingly smaller dimensions. Thus, throughout the 1990s, the industry’s response to the presumptive anomaly identified in the 1980s was a series of modular changes to the existing CMOS architecture (Henderson and Clark, 1990).

By the turn of the century, government investment in nanoelectronics concepts began to take off with the launch of the National Nanotechnology Initiative (NNI). Semiconductors are highly featured in pre-NNI agency reports (e.g. WTEC, 1998) and DARPA program managers funded “beyond CMOS” concepts in the early 2000s, building on DARPA’s 1990s ULTRA program.

The SRC’s primary research program continued to focus on continued scaling of CMOS. SRC’s research program in beyond CMOS technologies began in earnest with the establishment of two small programs in 1999 and 2000, Cross-Disciplinary Research and the Advanced Devices Thrust. An SRC report argued – optimistically – that “the physical and chemical understanding necessary to reach the ultimate limits of CMOS technology also will provide the basis for inventing new technologies that will eventually supplement CMOS” (Cavin, Herr, Zhirnov 2000). This dual mandate is mirrored by SRC’s funding for “ultimate” and “beyond” CMOS technologies. The modestly funded program focused on two parallel tracks: extending CMOS through non-classical structures and investigating novel device concepts (SRC, 2000a).

The 2001 International Technology Roadmap for Semiconductors included a new chapter – “Emerging Research Devices”. The structure of the chapter followed the same two-pronged approach: “non-classical CMOS” and “Emerging Logic” devices. The authors of the chapter noted that many of the devices were “speculative” and not directly competitive with CMOS (ITRS, 2001), but concluded with a rough quantitative comparison of emerging devices to CMOS.

The 2003 ITRS highlighted the ERD chapter’s discussions of post-CMOS devices as “pav[ing] the way to a complete technological revolution looming ahead towards the end of the next decade” (ITRS, 2003). The chapter went beyond a broad survey of emerging device concepts to include “a balanced, critical assessment of these emerging new device technologies” that “provides an industry perspective” (ITRS, 2003). The 2003 ERD chapter’s final section, “Emerging Technology – A Critical Review,” was built around a fundamental analysis of binary logic switching. In an ideal case of Constant’s (1980) presumptive anomaly, the analysis concluded, “even if entirely different electron transport devices are invented for digital logic, their scaling for density and performance may not go much beyond the ultimate limits obtainable with CMOS technology, due primarily to limits on heat removal capacity” (Zhirnov et al, 2003). The chapter concluded that none of the existing research devices were “viable emerging logic technologies for integration” (ITRS, 2003).
The SRC and industrial leadership responded to the lack of viable alternatives by engaging in a collaborative direction-setting process to identify new avenues for research. SRC’s “Novel Device Task Force” issued a “Research Needs for Novel Devices” document in May 2003 that summarized the situation:

“During the last several years there has been a huge amount of government-sponsored research loosely described as nanotechnology. This effort has lacked organization and direction and consequently is of limited immediate value to the semiconductor industry. However, the associated body of knowledge provides a rich resource that can be used as a starting point for targeted research programs” (SRC, 2003)

4.1.2 Industry-Government-University Response to Moore’s Wall

In July 2003, the SIA convened a subcommittee of senior industry executives to organize a “nanotechnology strategy” to leverage vast new public funds for nanotechnology research allocated by the NNI. Beginning in October 2003, SRC and NSF organized three industry-academia-government workshops, Silicon Nanoelectronics and Beyond (SNB), to discuss research directions for “Beyond CMOS” technologies with a particular focus on identifying promising alternative state variables. The workshops issued a set of 13 “research vectors” to guide future research efforts and established a host of offshoot groups to coordinate global research. By June 2004, the SIA board approved a program under the SRC, named the Nanoelectronics Research Initiative (NRI).
Table 3: Key participants in groups that shaped the NRI research program.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tak Ning</td>
<td>IBM</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Tom Theis</td>
<td>IBM</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Philip Wong</td>
<td>IBM</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>George Bourianoff</td>
<td>Intel</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Paolo Gargini</td>
<td>Intel</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Dimitri Nikonov</td>
<td>Intel</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bob Doering</td>
<td>TI</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Mihail Roco</td>
<td>NSF</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gernot Pomrenke</td>
<td>DoD</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>David Seller</td>
<td>NIST</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Ralph Cavin</td>
<td>SRC</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jim Hutchby</td>
<td>SRC</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daniel Herr</td>
<td>SRC</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Victor Zhinov</td>
<td>SRC</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimitri Antoniadis</td>
<td>MIT</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Mark Lundstrom</td>
<td>Purdue</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Kang Wang</td>
<td>UCLA</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Per the SIA Board, the NRI’s research objective was to “discover and reduce to practice via technology transfer to industry novel non-CMOS devices” by 2020 (SIA, 2004). To guide research toward this goal NRI adopted five of the thirteen vectors identified by the SNB workshops as the core of its program. In a white paper introducing the NRI, George Bourianoff (Intel) and Tom Theis (IBM) argued that the choice of research vectors would also directly shape the possible design space for NRI researchers: “the search for the next logic switch beyond CMOS will need to embrace some radical paradigm shifts and focus on key elements that limit conventional micro electronics” (i.e. limits to heat dissipation) (Bourianoff and Theis, 2004). NRI organizers described the program as “switch centric” research with the expectation of expanding the set of potential options available to industry and defining criteria for evaluating different options (Bourianoff, 2004). Bourianoff and Theis summarized the purpose of the selected research vectors:

“While this research agenda seems quite radical when stated as a coherent whole, in fact many elements of it already exist in the research community. Taken together, they provide the next level of description for what the ‘next logic switch beyond CMOS’ might look like. The discovery model of research nurtured by the NSF seems to have laid the foundations for this new technology in an uncorrelated, bottoms up, curiosity-driven fashion and our task now is to synthesize selected
elements into a coherent research plan. These selected elements would form the ‘goal’ of a research program planning process. After performing an assessment of the existing research inventory in university NSEC centers, MRSEC centers, MARCO centers and other centers of excellence, it will be possible to do a “gap analysis” of the requirements to reach the goal and create a plan to execute the gap analysis.” (Bourianoff and Theis, 2004) [emphasis original]

The industry’s intention of synthesizing existing research streams focused on research funded by the NSF and existing SRC programs. However, committee planning documents highlight the industry’s delicate balancing act with respect to existing government support for beyond CMOS research from NSF and DARPA. Absent from the above synthesis was DARPA-funded work in spintronics, molecular electronics, and other concepts dating back to the ULTRA programs of the 1990s. An NRI planning document describes the DARPA perspective as, “What are the new credible device concepts beyond CMOS that are not already being funded?” (SIA, 2004). Beginning in August 2004, industry representatives visited existing NSEC and MRSEC centers to identify overlap between existing research endeavors and the thirteen SNB workshop research vectors. Out of this process, NRI’s plan of execution took shape.

Figure 5 – Funding totals at DARPA and NSF for Beyond-CMOS technologies, 1991-2013. Source: NSF, 2013 and RDSS, 2013.
4.1.3 The NRI Model

Billing itself as a “goal-oriented, basic-science research program,” NRI averaged around $20 million in annual funding from its inception in 2005, a small figure in the context of ongoing federal funding in nanoelectronics research. Through 2006, SRC’s existing programs (GRC and FCRP) had focused on extension of the CMOS paradigm. Industry membership in these programs featured firms from several levels of the supply chain (e.g. equipment and materials suppliers) in addition to integrated device manufacturers. Additionally, these programs were majority-funded by industry members. Unlike previous SRC efforts, SIA expected NRI to be majority-funded by public sources. Further, unlike other SRC programs, all NRI member firms were integrated device manufacturers that both designed and manufactured their own products. These are the types of firms identified by Kapoor (2013) as being likely to contribute systemic innovations despite the industry’s specialized structure. Of the six firms, only IBM, the only diversified firm, continued to conduct a significant amount of basic research through the 1990s (Lim, 2004). The final structure and federal, state, and local government versus member-firm funding contributions can been seen in Figure 6.

Table 4: Summary of NRI Member firms for Fiscal Year 2006, in millions. Source: Standard and Poor’s, 2013 & IHS, 2006

<table>
<thead>
<tr>
<th>Firm</th>
<th>Net Income</th>
<th>Total Revenues</th>
<th>R&amp;D Expenses</th>
<th>R&amp;D Intensity</th>
<th>Global IC Sales Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMD</td>
<td>$ (166)</td>
<td>$ 5,649</td>
<td>$ 1,621</td>
<td>28.7%</td>
<td>8</td>
</tr>
<tr>
<td>Intel</td>
<td>$ 5,044</td>
<td>$ 35,382</td>
<td>$ 5,873</td>
<td>16.6%</td>
<td>1</td>
</tr>
<tr>
<td>IBM*</td>
<td>$ 9,492</td>
<td>$ 91,424</td>
<td>$ 5,682</td>
<td>6.2%</td>
<td>21</td>
</tr>
<tr>
<td>Micron</td>
<td>$ 408</td>
<td>$ 5,273</td>
<td>$ 656</td>
<td>12.4%</td>
<td>12</td>
</tr>
<tr>
<td>Freescale**</td>
<td>$(1,994)</td>
<td>$ 6,359</td>
<td>$ 1,200</td>
<td>18.9%</td>
<td>9</td>
</tr>
<tr>
<td>Texas Instruments</td>
<td>$ 4,341</td>
<td>$ 14,195</td>
<td>$ 2,195</td>
<td>15.5%</td>
<td>3</td>
</tr>
</tbody>
</table>

*IBM Microelectronics is only one portion of IBM’s overall portfolio of businesses **Freescale, formerly the semiconductor manufacturing division of Motorola, left NRI in 2008.

The NRI funded research through two primary mechanisms: device centric multi-university research centers and a set of projects co-funded with the NSF (see Figure 7). Impetus for the formation of the centers came from the firms themselves. Interviews with center directors indicated that firms approached the professors, encouraging them to organize a program around NRI’s goals. The first three centers were chosen without an open competition. A failed proposal to the NSF for a spintronics center led by a team from UCLA in coordination with Intel became the core of the first NRI center, WIN. TI and
IBM worked closely with universities and state governments in Texas and New York, respectively, to emulate the model put forth by Intel and the WIN center. In addition to funding from the consortium, each center received funding from state government, as well as start-up funding and equipment donations from the sponsoring firm. In California and Texas, established state-level programs also provided funding. IBM played an instrumental role in helping the MIND center get support from local and state governments.

NRI also began co-funding joint awards at NSF MRSECs and NSECs beginning in 2006. IBM’s Jeff Welser, director of NRI from 2006 through 2012, noted that NSF project research aimed to expand the scope of the NRI program. Industry executives stressed that the technology challenge facing the industry was highly uncertain and the NRI model offered a scope and breadth not possible in current corporate R&D labs. Tom Theis, explained:

“Well, even IBM with the resources we have here can't consider putting enough people on this problem. … Because of NRI, you've got 34 universities with top people involved thinking about this problem. IBM can't match that . . . And so, that [i.e., the larger university research] community is going to come up with ideas that IBM is not going to come up with. Now, IBM researchers can inject their own ideas into that mix. That's happening…, but for the amount of money that IBM puts into NRI, you could not do anywhere close to that amount of research within IBM” (Theis, 2013).
The formation of the NRI as the industry’s response to a presumptive anomaly shaped the processes by which it operated and offers insight into the potential for a public-private partnership to address technological needs through the development of new science. As noted above, the research vectors that formed the core of the NRI’s technical vision were designed to constrain the search space for new technologies (e.g. March, 1991; Levinthal and March, 1993; Powell and Gianella 2010). Several implications flowed from this decision. First, as Constant notes, “the scientific insight from which a presumptive anomaly is derived may and usually does contain simplifications or idealizations which render it not directly applicable to technological reality” (Constant, 1980 pg 17). Second, the selection process for technologies was shaped by both the complexity of technological systems and the traditions of practice of the technological community. The processes by which the NRI operated derived from the complexities of technological search. Based on interviews and participant observations, we identify several mechanisms by which the NRI engaged the semiconductor industry’s existing traditions of practice in its search for a new technological paradigm.

### 4.2.1 Validating Industry’s Need for Technology Outside the Prevailing Paradigm

Industry engagement in the NRI legitimized academic research into alternatives to CMOS, drawing greater numbers of researchers into the area. Due to the industry’s long history of failed alternatives (Choi and Mody, 2009), building momentum in research directions outside the existing paradigm played an important role in validating research outside the existing technological paradigm by researchers not only in but beyond those funded by NRI.

As a result, several industry researchers remarked that academic researchers initially responded with skepticism about NRI’s mission and doubted the sincerity of the industry’s quest for an entirely new device concept. Founding NRI director Jeff Welser summed up the disconnect between academic leaders and industry members:

“For the first couple years, our problem with them was they would not think far enough out. They were so used to being basically beat up by us saying, . . . [when] they'd come up with some idea . . ., ‘That's ridiculous, that will never work.’ Because that was the SRC program and the industry mindset had been very much on, ‘I need something for the next N plus two nodes or whatever, so don't tell me about this weird material.’ So, they didn't believe us. Even though we said, ‘We want something that's not a FET,’ everything they showed us was an FET. It was different materials, but everything was an FET. And, we were like, ‘But we said we don't want an FET.’ And they said, ‘Yeah, yeah, but we don't believe you.’” (Welser, 2013)
For industry researchers this disconnect threatened to derail the program’s timeline. At the 2012 annual review, several industry members spoke of the struggle of coordinating the work of academic researchers with industry’s needs. One industry liaison, who worked closely with academic researchers, said, “What we meant by 2020 . . . [the principal investigators] thought [was] ‘proof of concept by 2020.’ . . . [But] we meant commercialized devices” (Colombo, 2012). Academic researchers were surprised to find industry to be “more patient. As one academic noted, they “allowed us to work on some of the craziest ideas” (NRI Center Director, 2013). The industry’s openness to “crazy ideas” would validate the field and draw in other researchers. A federal agency representative working with the NRI stressed that once the industry had proven its commitment to the NRI, attendance at NRI’s annual and center reviews regularly began to include researchers from federal agencies and professors not involved with NRI (Federal Agency Researcher, 2013). Welser explained that non-NRI affiliated researchers used the program to justify their own work in the area,

“I’d be at other unrelated conferences, and then people would be talking about their own work, which we weren’t funding and had nothing do with us, and they were talking about the NRI and they’d say, ‘NRI, you know, this industry thing is working on this-this problem. This is one of the things they’ve identified as being really important. Here’s some work I’m doing, which actually is heading in helping that same problem.’”

Welser added that such legitimization provided “value that we had not anticipated because we had not seen that, I don’t think, with the other SRC programs as much in the past” (Welser, 2013).

4.2.2 Center Directors Link Basic Researchers to Industry

In his theory of technological revolution, Constant maintains that “because presumptive anomaly is science-based, it is most likely to be recognized initially, and in some cases, singularly, by those very close to the intellectual foundations of their respective fields” and often by “young outsiders” (Constant, 1980 pg 17). Yet, NRI center directors were chosen to be “typical engineering professors” who had a long history of interaction with industry research either through SRC or in industrial labs (Welser, 2013). Kang Wang and Sanjay Banerjee, directors of the WIN and SWAN centers, respectively, had each received more than 25 contracts from SRC prior to 2006. Banerjee had worked at Texas Instruments in the 1980s before entering academia. Founding director of the INDEX center, Alain Kaloyeros, had led the development of the Albany Nanotech Center and worked closely with IBM and other semiconductor firms since the mid-1990s. Alan Seabaugh, the director of the MIND center at Notre Dame, had worked at Texas Instruments and Raytheon Central Laboratories before moving to Notre Dame in 1999.
Center directors worked closely with NRI’s industry technical advisory group on shaping the focus of each center’s research program and building collaborative teams around new device concepts. NRI centers tapped a pool of researchers that included scientists, primarily from fields such as physics and materials science, who had not previously worked with the semiconductor industry. Collaborating with these “outsiders” introduced new challenges for industry personnel accustomed to working with engineering professors who better understood the industry’s technical challenges. NRI center directors played an important role in bridging the distinct academic and industrial research communities, acting as both “boundary spanners” (Fleming and Waguespack, 2007) and translators (Aitken, 1977). At the SWAN center, center leadership organized workshops on CMOS device operation to bring scientists unfamiliar with electrical engineering concepts up to speed on the challenge facing the industry. These seemingly simple interactions served to educate scientists on basic boundary conditions for identifying a successor device.

4.2.3 Developing a Common Language and Goals

Industry-academia interactions through NRI shaped the focus and direction of academic research through both formal and informal mechanisms. While most NRI researchers were scientists interested in understanding fundamental mechanisms to achieve proof of concept for new devices, industry attendees often focused on engineering challenges to designing systems around new computational variables. Industry participants found that constant industry-academia interactions through NRI helped the industry “define the problem” (Theis, 2013) and “give focus” to academic research (Industry Assignee #1, 2013). These responses suggest that one effect of NRI’s research program was a push by industry participants for academic researchers to adopt technology-driven search practices as opposed to science-driven search (Nightingale, 1998).

Evaluating the progress of NRI device concepts that used different computational state variables remained one of the program’s principal challenges. The NRI established a benchmarking program in 2000 that attempted to compare devices being developed by different NRI centers using common metrics such as switching delay, energy, and device area for realized circuits based on new device types (Bernstein et al., 2010; Nikonov and Young, 2013). Figure 8 below shows one of the initial benchmarking report’s main outputs.

The NRI benchmarking program induced academic researchers with backgrounds in fundamental science to consider system design and manufacturing limitations for new device concepts. NRI industry leadership instituted a series of workshops and lectures for students and professors to educate them on “what needs to be done to make devices interesting to industry” (Marshall, 2013). Although those
involved with the benchmarking caution that the accuracy of the results is limited because few devices have working prototypes, the process shaped both academic and industry research directions.

Current NRI Director Tom Theis argued that the “benchmarking program has really helped people in the academic community to understand what the problem is” because it allows an apples-to-apples comparison. Similarly, Dimitri Nikonov, an Intel liaison and a leader of the benchmarking program, said the program allowed the industry “to set future research directions” by focusing efforts on what looked most promising, akin to the search processes that Vincenti (1993) describes as selection and retention. While some academic participants worried initially that the benchmarking program would be used to kill projects, afterward they admitted that they had gained “a better understanding of the problems facing the industry,” and they subsequently reshaped their device concepts to overcome initial limitations (Theis, 2013).

Figure 8 - NRI benchmarking results comparing switching speed and energy for NRI device concepts. Each labeled dot represents a new device type investigated by NRI researchers. Source: Bernstein et al., 2010

Unlike market-mediated technology selection (i.e. “eras of ferment”) or traditional bottoms-up approaches to funding of scientific research, the NRI’s benchmarking was a formal program mediated by employees of existing companies. While the implications of this structure on the industry’s future technology directions remained to be seen, the benchmarking actively shaped academic and industrial research, focusing researchers’ work on concepts most promising according to a select set of metrics. The consortium also offered a collaborative space for academic and industrial researchers not only to develop and agree upon these metrics but also to discuss the possibilities and limitations of new devices beyond the metrics proposed.
The continued use of the NRI benchmarking exercise, which Kaplan and Tripsas have (in a different context) described as “technological framing” (Kaplan and Tripsas, 2008), reflected the industry’s desire to maintain its pre-existing business model and associated core competences. Or, as Dosi (also in a different context) has captured the process, NRI’s benchmarking objective served to “extend the historical cost and performance trends for information technology” (Welser, 2007) even while considering new technology paradigms (Dosi, 1982). It is impossible to say whether this focus and the associated orchestration of technology development by NRI’s large, incumbent firms was necessary for maneuvering such a complex, vertically-disintegrated, technology- and capital-intensive industry through extreme a technological discontinuity; alternatively, the focus might have negatively limited the technological possibilities of the program, as well as more broadly.

4.2.4 Industry Research Engagement and Technology Transfer

4.2.4.1. Assignees and Liaisons Influence Academic and Industry Research Directions

Industrial researchers engaged with NRI academic research through a variety of mechanisms. Two governing boards, the Technical Program Group and Governing Council, were comprised of researchers and research managers from industry and government. They formed the core group of technical reviewers who evaluated the merits of different proposals as well as offering input on management issues throughout the duration of the program. Industry researchers also engaged directly with academic research as both industry liaisons and assignees. NRI assignees were industry researchers designated by member firms to work on-site at academic research centers in coordination with NRI researchers. Member firms received a discount toward their annual NRI dues for placing assignees at NRI centers.

NRI’s industry assignees interacted with NRI centers’ researchers differently based on the size, capabilities, and technological focus of their employing firm. The smaller firms, Micron and Global Foundries, did not have strong in-house research programs to complement NRI’s research. Each firm had only one assignee, and both assignees described their primary function as learning and studying NRI device concepts. By contrast, the larger firms (Intel, IBM, and TI) each sponsored a regional research center where their assignees played a hands-on role in shaping NRI’s research through a variety of roles.

In interviews, academic researchers noted that while joint work was published, the industry researchers functioned primarily as employees of their firms. One Intel assignee worked with the WIN center at UCLA for 5 years, where he described his role as “getting people to collaborate and integrating all spin-related research projects” (Industry Assignee #1, 2013). TI’s assignee worked primarily with the SWAN center in Texas and co-patented six inventions with SWAN researchers, nearly half of the NRI’s total patent output. Both assignees noted that they limited their own patenting in order to manage relationships
with academic researchers and stressed that they were most effective as mentors for projects and students. The ongoing feedback industry researchers provided to participants about research directions also influenced research trajectories. Some academic researchers occasionally bristled at the loss of research freedom, and this was a source of tension early on in the program. In interviews, academic researchers who had accepted NRI funds were coy about discussing the demands placed on them by NRI’s funding.

NRI assignees and liaisons constituted a special group of individuals who had interacted with each other through their firms’ participation in other SRC and ITRS working groups. For example, several of the NRI liaisons and assignees also contributed to the ITRS ERD chapter and shared NRI research findings through that publication. In December 2013, NRI’s industry representatives presented NRI benchmarking results at the industry’s largest device conference, IEDM. At the same conference, NRI industry representatives presented a short-course on “Beyond-CMOS” devices and shared a paper summarizing the results of NRI’s benchmarking study. NRI leadership noted that researchers working on neuromorphic devices and superconducting circuits – devices not in NRI’s research portfolio – mentioned that they plan to adopt benchmarking metrics similar to NRI’s process for their own devices (Nikonov, 2013).

The many roles played NRI’s assignees and liaisons worked to shape research directions differed markedly from those played by industry employees in short-term-focused consortia (e.g. SEMATECH, MCC, VLSI), where employees primarily worked on managing and evaluating technical progress of development projects (Grindley et al., 1994). Their roles were also distinct from university-industry collaborations on individual projects where the focus was on bringing ideas back into the firm. In part because NRI’s early research results were less directly appropriable by their employing firms, the assignees focused on improving research directions more broadly.

4.2.4.2. Intellectual Property Challenges Given a Highly Uncertain Technical Future

NRI also allows us to observe phenomena associated with consortia as research exits the pre-competitive stage and enters the competitive stage. Unlike previous industry consortia NRI offers an example of coopetition where consortium members cooperate on the research stage but compete in the product stage. In this situation, knowledge spillovers may erode competitive advantage in the new technology. Within this context, the role of patent rights and the nature of technology transfer from academia to industry labs became potential sources of friction.

The lack of clearly prescribed roles and rights of industrial assignees – e.g., were they allowed to conduct their own experiments at NRI centers? – led to disagreements over hypothetical patent rights. In at least two cases, disagreements between universities and the NRI consortium delayed the arrival of industrial assignees and forced a restructuring of their eventual roles within the centers. These instances
showed that despite SRC’s decades of experience and cooperation with universities on intellectual property rights and procedures, the high degree of uncertainty of the NRI project and the potential value of a fundamental patent from NRI-funded research inevitably made this element of the public-private partnership both fragile and susceptible to dispute.

The MIND center at Notre Dame primarily conducted research on the tunneling FET, and based in part on work done through the NRI, Intel began working internally on a tunneling FET. Through 2013, the MIND center explored a variety of device configurations (i.e. material stacks and device geometries) for the tunnel FET. When Notre Dame proposed a new center, housed in the STARnet program, based on the core of the work from the MIND center, Intel blocked work on specific configurations and materials systems for the tunneling FET that had been pursued by MIND. Intel researchers presented work on tunnel FETs at industry conferences, most notably at IEDM 2013, that utilized one of the material stacks previously pursued within MIND. This transition from pre-competitive to competitive research also reflects the broader flow of information between academic researchers and industry liaisons within the program. Several academic interviewees noted that the process by which firms take research internally is opaque.

4.2.5. Developing Human Capital for a New Technology Paradigm

For over 30 years, SRC’s traditional research programs trained students in technologies immediately relevant to member firms’ development programs. Industry executives estimated that this enhancement of human capital through SRC-funded research saved their firms roughly $100,000 per student in training costs. NRI-funded students, however, did not conduct research in areas immediately relevant to the industry’s development programs. Unlike other SRC programs where 70% of graduates went to work at SRC member firms, only 50% of NRI students entered industry. Instead, industry executives argued that NRI was cultivating a new network of multi-disciplinary technologists trained to address the largest, unprecedented research needs of the semiconductor industry. One industry executive described research participants’ human capital advantage this way:

“Every student that has been brought into the NRI program has had some experience with thinking outside of what I call ‘the FET box,’ . . . and that's a liberating thing that they're going to remember that for the rest of their career. They're going to remember that there's other ways to compute besides this device. (Theis, 2013)

One former student, now a researcher at an NRI member firm and an NRI industry assignee, acknowledged that NRI’s approach encouraged more cross-disciplinary collaboration, thus training him
better for work in a corporate research environment. This researcher highlighted his experience as a graphene theorist working with graphene experimentalists at a different university affiliated with the same NRI center (Industry Assignee #2, 2013).

Bringing NRI students’ knowledge to bear on industrial problems, however, posed challenges. One industrial assignee working onsite with NRI-funded physicists noted that the graduate students he collaborated with were “much more interested in fundamental physics than realistic problems” (Industry Assignee #3, 2013). Additionally, professors involved with the NRI showed caution about their students’ employment prospects in the industry. One center director noted that he was nervous because, “I had them working on these crazy ideas, but the [industry] jobs are in CMOS” (NRI Center Director, 2014). An NRI-affiliated professor added that although his graduate students were trained exclusively in non-silicon CMOS devices, the department had decided to maintain its undergraduate curriculum’s focus on CMOS for the foreseeable future because that was where students would find jobs (NRI Academic Researcher, 2014).

Unlike collaborations focused on more incremental technology change, NRI’s industry members—and, to the extent of the program’s open publication policies, the semiconductor industry more broadly—benefited mostly from strengthening the broader research ecosystem. Graduates that take positions in academia or government agencies may influence collaborators or shape funding decisions based on the knowledge they gained in NRI-funded centers. These ancillary benefits may help create the understanding and capability necessary throughout the technical community to change the existing paradigm. At the same time, these graduate students may struggle to find jobs in firms now focused almost entirely on achieving the last few generations of CMOS-based integrated circuits. Graduate students who do not go to or eventually leave incumbent firms could also become sources of change in the existing paradigm outside the incumbent firms in the industry.

6. Discussion: Is Creative Destruction Necessary?

For over 70 years, scholars of technology change have built upon Schumpeter’s notion of “creative destruction,” whereby innovative entry by entrepreneurs is the force that sustains long-term economic growth, even as it destroys the value of established companies (Schumpeter, 1942). Subsequent scholarly work has focused on why and when new technologies lead to the destruction of established firms. Abernathy and Clark (1985) argue that established firms will only be destroyed when the incoming technological innovation requires new core competencies not held by the established firms. Henderson and Clark build on this work, arguing that architectural innovations (in contrast to innovations in discrete modules that do not change the overarching product architecture) require new core competencies not held
by established firms. Chesbrough and Kusunoki (2001) add the concept of a modularity trap, in which incumbent firms lack incentives to maintain knowledge about how to innovate on the larger system architecture in a vertically disintegrated industry where each firm contributes to innovation on a singular modular part. In contrast to this focus on core competencies, Christensen argues that established firms will be destroyed by new technologies when the new technologies shift the performance attributes important to customers (Christensen 1997).

The technological discontinuity facing the semiconductor industry clearly requires new core competencies as well as architectural innovation of established firms throughout the supply chain – from integrated device manufacturers to equipment and material suppliers to software programmers themselves. Interestingly, however, Constant points out that in the case of aerospace industry, despite dramatic changes in technologies, materials, and the architecture of engines as well as planes themselves, “In addition to technological heritage, there was a remarkable persistence of the organizations and, in some cases, of the same men in the steam turbine, the first gas turbines, and finally the jet engine” (Constant, 1980 pg 18). The historical case of the turbojet raises questions as to whether the Schumpeterian creative destruction is truly a necessary precondition for radical, architectural, or disruptive innovation across a technological discontinuity. Indeed, for the case of semiconductors, successful innovation across this technological discontinuity without the engagement of the established firms is difficult to imagine.

7. Conclusion

The semiconductor industry faces an unprecedented technical challenge: to extend “Moore’s Law” it must develop a novel computing device for which the underlying science is unknown and that will require vast changes in the industry’s entire supply chain. The case of NRI offers insights into the formation and subsequent execution of a public-private partnership designed to address a technological grand challenge. Our findings simultaneously shed light on potentially successful methods and on challenges facing public-private partnerships (and collaborative R&D more broadly) in addressing technological discontinuities where there is great uncertainty and the underlying science to support the solution is unknown. We show how key individuals within the technical community, relying upon existing institutions eventually converged on the technical problem at hand and an institutional response, the Nanoelectronics Research Initiative (NRI). NRI’s research program represented a willingness by a small segment of the industry to depart from the industry’s prevailing technology paradigm to establish a new technology paradigm while maintaining its existing performance trajectory (Dosi, 1982). These firms’ willingness to support long-term research through the NRI and the NRI’s ability to leverage
federal, state, and even local research funding have helped to build momentum in non-CMOS FET switch research among academics in several regions of the U.S.

While this study offers important initial insights about public-private partnerships for long-term technology development, there are limits to its generalizability. While other industries may face similar technological grand challenges, NRI is a byproduct of the unique institutional ecology of the semiconductor industry. U.S. semiconductor firms have over 30 years of experience with cooperative R&D. Additionally, the industry is noted for the limited role patents play in appropriating returns from R&D (Levin et al., 1987; Hall and Ziedonis, 2001). At the same time, many aspects of the problem facing the semiconductor industry are common to other industries including vertical disintegration, difficulty appropriating the returns to innovations, and uncertainties about future technology directions (Powell and Giannella 2010). Likewise, multiple aspects of the semiconductor industry case resonate in other contexts, including extensive reliance on inter-organizational collaboration in R&D (e.g. biotechnology, see Powell and Giannella 2010); the technology hierarchy with high level of complementarities requiring system-wide and architectural innovation to solve the problem (e.g. aerospace); and the extraordinarily high capital costs of both equipment and R&D (e.g. pharmaceuticals and chemicals).

Changes in the scale and structure of corporate research and development over the last three decades have made it harder for firms to respond to long-term technology grand challenges. The persistence of the same organizations (and in some cases individuals) in the steam turbine, the first gas turbines, and finally the jet engine, raises questions as to whether the type of Schumpeterian creative destruction assumed by scholars and the popular press alike is truly a necessary precondition for radical, architectural innovation across a technological discontinuity such as that of scaling Moore’s Wall. Indeed, for the case of semiconductors it is unclear whether innovation through the technological discontinuity without engagement of the established firms is possible. Whether the NRI and complementary institutions across the nation are sufficient to address the challenges currently facing the semiconductor industry and the social and economic growth associated with scaling Moore’s Wall remains to be seen.

Acknowledgments

Thanks to NIST and NSF Science of Science Policy for their generous funding of this research. Also, thanks to the Semiconductor Research Corporation and the many industry, government, and academia participants who have been patient with their time and resources and made this study possible. All views expressed are those of the authors.
References


**Referenced Data Sources:**


Bourianoff, G., 2006. NSEC visit report and NRI research vectors SNB II. Presentation to SIA Technology Strategy Committee. December 15, 2005. Received directly from Presenter.


Interview via phone with NRI Center Director on June 26, 2013.

Interview via phone with NRI Center Director on February 2, 2014.

Interview with Federal Agency Researcher on September 13, 2013.

Interview with Industry Assignee #1 in person at NRI Annual Review on October 22, 2013.

Interview with Industry Assignee #2 in person at IBM TJ Watson Lab on August 16, 2013.

Interview with Industry Assignee #3 in person at GLOBALFOUNDRIES on August 5, 2013.

Interview with NRI Academic Researcher via Skype on April 3, 2014.


