

ISBN 978 94 6159 511 9

© copyright Giorgio Triulzi, Maastricht 2015

Printing: Datawyse / Universitaire Pers Maastricht

Cover illustration by Juan Sebastian Taborda (for inquiry: [sebtav19901@gmail.com](mailto:sebtav19901@gmail.com))

# LOOKING FOR THE RIGHT PATH

## Technology Dynamics, Inventive Strategies and Catching-up in the Semiconductor Industry

DISSERTATION

to obtain the degree of Doctor at Maastricht University,  
on the authority of the Rector Magnificus Prof. dr. L.L.G. Soete,  
in accordance with the decision of the Board of Deans,  
to be defended in public on Wednesday 16<sup>th</sup> December 2015, at 09:45 hours

by

**Giorgio Triulzi**

**Supervisor**

Prof. Dr. Bart Verspagen

**Co-supervisor**

Dr. Önder Nomaler

**Assessment Committee**

Prof. Dr. Robin Cowan (Chair)

Prof. Dr. Koen Frenken (Utrecht University)

Prof. Dr. Franco Malerba (Bocconi University)

Dr. Lili Wang

*To Jennifer, my parents and my sister*





# CONTENTS

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>INTRODUCTION: COMPLEXITY, TECHNOLOGICAL CHANGE AND CATCHING UP .....</b>   | <b>1</b>  |
| 1.1      | CONCEPTUAL FRAMEWORK AND STRUCTURE OF THE THESIS .....  | 5         |
| <b>2</b> | <b>TECHNOLOGICAL AND INDUSTRIAL EVOLUTION OF THE GLOBAL SEMICONDUCTOR INDUSTRY.....</b>                                 | <b>9</b>  |
| 2.1      | INTRODUCTION .....  | 10        |
| 2.2      | TECHNOLOGICAL CHANGE, INDUSTRIAL DYNAMICS AND CATCHING-UP.....  | 10        |
| 2.3      | ARE SUCCESSFUL CATCHING-UP COUNTRIES SPECIAL?.....  | 15        |
| 2.3.1    | <i>Similarities and differences of initial conditions .....</i>   | <i>15</i> |
| 2.3.2    | <i>Sources of technical improvements.....</i>   | <i>17</i> |
| 2.3.3    | <i>Speed of technical improvements .....</i>  | <i>19</i> |
| 2.4      | CONCLUSIONS .....   | 20        |
| <b>3</b> | <b>ENGINEERING TRAJECTORIES, RANKING OF DESIGN PROBLEMS AND CATCHING-UP .....</b>                                       | <b>23</b> |
| 3.1      | INTRODUCTION .....  | 24        |
| 3.2      | THEORETICAL FRAMEWORK.....  | 25        |
| 3.3      | METHODOLOGY.....  | 30        |
| 3.3.1    | <i>Identification of technically influent patents and of the main paths of technical solutions ..</i>                   | <i>30</i> |
| 3.3.2    | <i>Measuring the ranking of the system of engineering problems and the stability of problem-solving approaches.....</i> | <i>32</i> |
| 3.3.3    | <i>Measuring centrality on the paths of engineering improvements .....</i>  | <i>33</i> |
| 3.3.4    | <i>Measuring path-following, path-changing, ranking-following and ranking-changing propensity .....</i>                 | <i>35</i> |
| 3.4      | DATA.....   | 38        |
| 3.4.1    | <i>Descriptive statistics of the Network of Main Paths.....</i>   | <i>38</i> |
| 3.5      | FINDINGS .....  | 40        |
| 3.5.1    | <i>Direction of technological change.....</i>   | <i>40</i> |
| 3.5.2    | <i>Players' inventive strategies.....</i>   | <i>48</i> |
| 3.5.3    | <i>Effectiveness of players' strategies.....</i>  | <i>51</i> |
| 3.6      | CONCLUSIONS .....   | 52        |
| <b>4</b> | <b>LIFE CYCLE OF TECHNOLOGY DOMAINS AND COMPARATIVE TECHNOLOGICAL ADVANTAGE.....</b>                                    | <b>55</b> |
| 4.1      | INTRODUCTION .....  | 56        |
| 4.2      | TECHNOLOGY AND INDUSTRIAL DYNAMICS OF THE GLOBAL SEMICONDUCTOR INDUSTRY .....   | 58        |
| 4.3      | THEORETICAL FRAMEWORK .....   | 62        |
| 4.4      | METHODOLOGY AND PRELIMINARY DATA ANALYSIS.....  | 66        |
| 4.4.1    | <i>Identification of technologically influent patents.....</i>  | <i>67</i> |
| 4.4.2    | <i>Grouping technologically influent patents into technology domains .....</i>  | <i>69</i> |
| 4.4.3    | <i>Characterizing technology domains according to their life cycle stage.....</i>                                       | <i>70</i> |
| 4.4.4    | <i>Measuring comparative technological advantage along the life-cycle .....</i>   | <i>76</i> |
| 4.5      | FINDINGS .....  | 78        |
| 4.5.1    | <i>New and incumbent innovators' revealed technological advantage .....</i>   | <i>79</i> |

|          |   |            |
|----------|---|------------|
| 4.5.2    | <i>Countries' revealed technological advantage</i> .....  | 81         |
| 4.6      | DISCUSSION AND CONCLUSIONS.....   | 85         |
| <b>5</b> | <b>DO I FIT? INNOVATIVE SURVIVAL, TECHNOLOGY DYNAMICS AND THE EVOLUTION OF KNOWLEDGE</b> .....  | <b>87</b>  |
| 5.1      | INTRODUCTION .....  | 88         |
| 5.2      | THEORETICAL FRAMEWORK AND RESEARCH HYPOTHESES.....  | 90         |
| 5.3      | DATA PREPARATION AND DESCRIPTION.....   | 95         |
| 5.3.1    | <i>Identification of influential patents and core innovators</i> .....  | 95         |
| 5.3.2    | <i>Identification of Semiconductor technology domains</i> .....   | 100        |
| 5.4      | MODEL.....  | 101        |
| 5.5      | MEASUREMENT OF THE KEY VARIABLES .....  | 103        |
| 5.5.1    | <i>Knowledge persistence</i> .....  | 104        |
| 5.5.2    | <i>Firms' knowledge breadth and depth</i> .....   | 107        |
| 5.5.3    | <i>Measures of knowledge modularity</i> .....   | 109        |
| 5.5.4    | <i>Control variables</i> .....  | 110        |
| 5.6      | TRENDS OF KNOWLEDGE MODULARITY AND PERSISTENCE.....   | 110        |
| 5.7      | ECONOMETRIC ISSUES .....  | 111        |
| 5.8      | FINDINGS .....  | 116        |
| 5.9      | CONCLUSIONS.....  | 128        |
| <b>6</b> | <b>CONCLUSIONS</b> .....  | <b>131</b> |
| 6.1      | SUMMARY OF THE THESIS' CONTRIBUTIONS AND FINDINGS.....  | 131        |
| 6.2      | IMPLICATIONS OF THIS PHD DISSERTATION FOR TECHNOLOGY AND INNOVATION STUDIES .....   | 133        |
| 6.3      | POLICY AND MANAGERIAL IMPLICATIONS.....   | 136        |
| 6.4      | RESEARCH AGENDA.....  | 138        |
| 6.4.1    | <i>Causes</i> .....   | 138        |
| 6.4.2    | <i>Effects</i> .....  | 138        |
| 6.4.3    | <i>Predictions</i> .....  | 139        |
| <b>7</b> | <b>REFERENCES</b> .....   | <b>141</b> |
| <b>8</b> | <b>APPENDICES</b> .....   | <b>149</b> |
|          | <b>APPENDIX TO CHAPTER 2</b> .....  | <b>150</b> |
| A.2.1    | NORMALIZATION OF THE NUMBER OF BETWEEN-FIRMS, WITHIN-COUNTRY CITATIONS BY<br>RANDOMIZING CITATION NETWORKS.....                       | 150        |
| A.2.2    | EMPIRICAL CUMULATIVE DISTRIBUTION FUNCTIONS OF THE PROBABILITY OF CITING A PATENT AS<br>A FUNCTION OF ITS AGE BY CITING COUNTRY ..... | 151        |
|          | <b>APPENDIX TO CHAPTER 3</b> .....  | <b>152</b> |
| A.3.1    | VISUAL ANALYSIS OF THE MAIN COMPONENT OF THE NETWORK OF MAIN PATHS.....   | 152        |
| A.3.2    | STATISTICAL ANALYSIS OF THE RELATIONSHIP BETWEEN PFI, POP, PATHC AND PATENT COUNT ....  | 156        |
|          | <b>APPENDIX TO CHAPTER 4</b> .....  | <b>157</b> |
| A.4.1    | ALTERNATIVE BEGINNING AND END OF THE ARCHETYPAL LIFE-CYCLE OF TECHNOLOGY DOMAINS....  | 157        |
| A.4.2    | PLOTS OF THE MAIN COMPONENT OF THE NETWORK OF MAIN PATHS (NMPs) .....   | 158        |
| A.4.3    | NEWMAN'S COMMUNITY DETECTION ALGORITHM .....  | 161        |

|                                   |  |            |
|-----------------------------------|--|------------|
| A.4.4                             | ANDERSON-DARLING TEST RESULTS.....   | 162        |
| A.4.5                             | TOPIC ANALYSIS OF THE MAIN TECHNOLOGY DOMAINS OF THE SEMICONDUCTOR INDUSTRY<br>BETWEEN 2001 AND 2006 .....                                     | 163        |
| A.4.6                             | SRTA TABLES AT THE FIRM LEVEL.....   | 168        |
| <b>APPENDIX TO CHAPTER 5.....</b> |  | <b>173</b> |
| A.5.1                             | RANKING OF OBSERVATIONS BY KNOWLEDGE DEPTH .....   | 173        |
| A.5.2                             | EMPIRICAL CONDITIONAL PROBABILITIES OF SURVIVAL.....   | 174        |
| A.5.3                             | KOLMOGOROV-SMIRNOF, ANDERSON-DARLING AND STUDENT'S T TESTS OF THE ASSOCIATIONS<br>OF KNOWLEDGE BREADTH AND DEPTH WITH INNOVATIVE SURVIVAL..... | 175        |
| A.5.4                             | REGRESSION RESULTS USING PROBIT MODELS .....   | 178        |
| A.5.5                             | REGRESSION RESULTS USING COMPLEMENTARY LOG-LOG MODELS.....   | 181        |
| A.5.6                             | SCATTERPLOT OF CORE INNOVATORS IN THE KB-KD SPACE .....  | 184        |
| A.5.7                             | THE LINKAGE BETWEEN MICRO- AND MACRO-DIVERSIFICATION STRATEGIES .....  | 185        |
| <b>9</b>                          | <b>VALORISATION .....</b>  | <b>189</b> |
| 9.1                               | DEGREE OF INNOVATIVENESS .....   | 189        |
| 9.2                               | SOCIO-ECONOMIC RELEVANCE .....   | 190        |
| <b>10</b>                         | <b>ABOUT THE AUTHOR.....</b>   | <b>192</b> |

# LIST OF TABLES

|  |     |
|--|-----|
| TABLE 1: METHODOLOGICAL FRAMEWORK ROADMAP .....  | 30  |
| TABLE 2: COMPARISON OF NETWORKS' SIZES.....  | 39  |
| TABLE 3: RELATIVE SIZE OF THE LARGEST NMPs' COMPONENT .....  | 39  |
| TABLE 4: PATENT PERSISTENCE WITHIN THE MAIN COMPONENT OF THE NETWORK OF MAIN PATHS.....  | 41  |
| TABLE 5: LIST OF PATH-CHANGING PATENTS.....  | 45  |
| TABLE 6: TOP PATH-FOLLOWER PATENTS.....  | 46  |
| TABLE 7: METHODOLOGICAL ROADMAP TO ANALYSE THE LIFE CYCLE OF TECHNOLOGY DOMAINS AND<br>SPECIALIZATION PATTERNS OF NEW AND INCUMBENT INNOVATORS ..... | 67  |
| TABLE 8: BASIC NETWORK STATISTICS.....   | 69  |
| TABLE 9: BASIC STATISTICS FOR THE TECHNOLOGY DOMAINS IDENTIFIED BY NEWMAN'S ALGORITHM .....  | 70  |
| TABLE 10: PATENT DISTRIBUTION QUANTILE BORDERS BY PATENT TYPE AND LIFE CYCLE STAGE .....   | 75  |
| TABLE 11: MOVEMENTS FROM ONE LIFE CYCLE STAGE TO THE OTHERS OVER CONSECUTIVE PERIODS.....  | 76  |
| TABLE 12: NUMBER OF FIRMS BY GEOGRAPHIC ORIGIN AND CATEGORY .....  | 79  |
| TABLE 13: METHODOLOGY ROADMAP .....  | 94  |
| TABLE 14: SIZE OF THE NMPs SAMPLE AND CATEGORIES OF INNOVATORS.....  | 98  |
| TABLE 15: CORE INNOVATORS' FUNCTIONAL AND GEOGRAPHIC COMPOSITION .....   | 99  |
| TABLE 16: BASIC STATISTICS FOR THE IDENTIFIED SEMICONDUCTOR TECHNOLOGY DOMAINS.....  | 101 |
| TABLE 17: MEASURING KNOWLEDGE GENETIC HERITAGE ACROSS PATENTS.....   | 106 |
| TABLE 18: MEASURING KNOWLEDGE PERSISTENCE .....  | 107 |
| TABLE 19: MEASURING FIRM-TECHNOLOGY DOMAIN PROXIMITY.....  | 108 |
| TABLE 20: KNOWLEDGE TAXONOMY OF CORE INNOVATORS .....  | 109 |
| TABLE 21: TRENDS OF KNOWLEDGE MODULARITY AND PERSISTENCE .....   | 111 |
| TABLE 22: CUTOFF VALUES OF KB CAT, KNOWCLASS4 AND KNOWCLASS6.....  | 114 |
| TABLE 23: CORRELATION BETWEEN INDEPENDENT VARIABLES AND COLLINEARITY DIAGNOSTIC .....  | 114 |
| TABLE 24: POOLED AND PANEL DATA LOGIT MODEL ESTIMATION .....   | 119 |
| TABLE 25: POOLED AND PANEL DATA LOGIT MODEL ESTIMATION WITH SPLIT SAMPLE.....  | 121 |
| TABLE 26: POOLED DATA LOGIT MODELS ESTIMATION WITH KNOWCLASS.....  | 122 |
| TABLE 27: PANEL DATA RANDOM EFFECT LOGIT MODELS ESTIMATION WITH KNOWCLASS.....   | 123 |
| TABLE 28: AVERAGE MARGINAL EFFECTS (AMES) OF DIVERSIFICATION CONDITIONAL ON THE LEVEL OF<br>KNOWLEDGE PERSISTENCE .....                              | 127 |
| TABLE 29: LEGEND OF NODES' COLOUR, SHAPE AND SIZE .....  | 152 |
| TABLE 30: TOPIC ANALYSIS OF THE MAIN TECHNOLOGY DOMAINS OF THE<br>SEMICONDUCTOR INDUSTRY BETWEEN 2001 AND 2006 .....                                 | 164 |
| TABLE 31: SRTA FOR THE TOP TAIWANESE, KOREAN AND SINGAPOREAN FIRMS (1991-1995) .....   | 168 |
| TABLE 32: SRTA FOR THE TOP US AND JAPANESE PLAYERS (1991-1995).....  | 168 |
| TABLE 33: SRTA FOR THE TOP TAIWANESE, KOREAN AND SINGAPOREAN PLAYERS (1996-2000) .....   | 169 |
| TABLE 34: SRTA FOR THE TOP US AND JAPANESE PLAYERS (1996-2000).....  | 169 |
| TABLE 35: SRTA FOR THE TOP TAIWANESE, KOREAN AND SINGAPOREAN PLAYERS<br>(2001-2006 - MAIN COMPONENT OF THE NETWORK OF MAIN PATHS).....               | 170 |
| TABLE 36: SRTA FOR THE TOP US AND JAPANESE PLAYERS<br>(2001-2006 - MAIN COMPONENT OF THE NETWORK OF MAIN PATHS) .....                                | 170 |
| TABLE 37: SRTA FOR THE TOP TAIWANESE, KOREAN AND SINGAPOREAN PLAYERS<br>(2001-2006 - SECOND COMPONENT OF THE NETWORK OF MAIN PATHS) .....            | 171 |
| TABLE 38: SRTA FOR THE TOP US AND JAPANESE PLAYERS<br>(2001-2006 - SECOND COMPONENT OF THE NETWORK OF MAIN PATHS).....                               | 171 |

|  |     |
|--|-----|
| TABLE 39: TOP 10% OBSERVATIONS BY LN(KD) .....   | 173 |
| TABLE 40: STATISTICAL TESTS FOR DIFFERENCES IN THE EMPIRICAL CUMULATIVE DISTRIBUTIONS OF<br>KNOWLEDGE BREADTH AND DEPTH BETWEEN SURVIVAL AND EXIT GROUPS ..... | 178 |
| TABLE 41: POOLED AND PANEL DATA PROBIT MODEL ESTIMATION.....   | 178 |
| TABLE 42: POOLED AND PANEL DATA PROBIT MODEL ESTIMATION WITH SPLIT SAMPLE.....   | 179 |
| TABLE 43: POOLED AND PANEL DATA COMPLEMENTARY LOG-LOG MODEL ESTIMATION .....   | 181 |
| TABLE 44: POOLED AND PANEL DATA COMPLEMENTARY LOG-LOG MODEL ESTIMATION WITH SPLIT<br>SAMPLE.....   | 182 |

## LIST OF FIGURES

|   |    |
|---|----|
| FIGURE 1: THE GLOBALIZATION OF SEMICONDUCTOR INVENTIVE ACTIVITIES .....   | 11 |
| FIGURE 2: USPTO SEMICONDUCTOR PATENT SHARES BY COUNTRY OF INVENTION .....   | 12 |
| FIGURE 3: SHARE OF SEMICONDUCTOR SALES BY GEOGRAPHIC DESTINATION.....   | 12 |
| FIGURE 4: SHARE OF SEMICONDUCTOR SALES BY GEOGRAPHIC ORIGIN .....   | 12 |
| FIGURE 5: SEMICONDUCTOR PATENT SHARES OVER COUNTRY'S TOTAL PATENT COUNT .....   | 16 |
| FIGURE 6: NUMBER OF ORGANIZATIONS ACTIVE IN SEMICONDUCTOR PATENTING BY COUNTRY .....  | 16 |
| FIGURE 7: THE INCREASING RELIANCE ON NATIONAL SOURCES OF TECHNICAL IMPROVEMENTS .....   | 18 |
| FIGURE 8: INDUSTRY-BUILDING BY IMPROVING UPON LOCAL TECHNOLOGIES .....  | 19 |
| FIGURE 9: FAST IMPROVEMENT OF SEMICONDUCTOR TECHNOLOGIES BY SUCCESSFUL CATCHING-UP<br>COUNTRIES .....   | 20 |
| FIGURE 10: THE RELATIONSHIP BETWEEN TECHNOLOGICAL TRAJECTORIES, RANKING OF<br>DESIGN PROBLEMS AND ENGINEERING DESIGN TRAJECTORIES.....  | 27 |
| FIGURE 11: A DUMMY CITATION NETWORK .....   | 31 |
| FIGURE 12: ATTRACTIVENESS OF THE TWO LARGEST COMPONENTS OF THE NMPs.....  | 40 |
| FIGURE 13: PATENT PERSISTENCE WITHIN AND BETWEEN COMPONENTS.....  | 42 |
| FIGURE 14: IDENTIFYING SIGNIFICANT PATH-CHANGER AND PATH-FOLLOWER PATENTS.....  | 44 |
| FIGURE 15: IDENTIFYING PATH-CHANGING AND PATH-FOLLOWING FIRMS.....  | 49 |
| FIGURE 16: RANKING-CHANGING AND RANKING-FOLLOWING FIRMS.....  | 52 |
| FIGURE 17: MOST CENTRAL PLAYERS .....   | 52 |
| FIGURE 18: INDUSTRY TECHNOLOGY DYNAMICS.....  | 61 |
| FIGURE 19: NEW ENTRANT INNOVATORS BY COUNTRY OF ORIGIN .....  | 62 |
| FIGURE 20: NESTED HIERARCHY OF LIFE CYCLES.....   | 63 |
| FIGURE 21: ARCHETYPAL LIFE-CYCLE OF A GIVEN TECHNOLOGY DOMAIN STARTING WITH A<br>BREAKTHROUGH .....   | 65 |
| FIGURE 22: THE RELATIONSHIP BETWEEN PERSISTENT OLD PATENT, NEW OLD PATENTS<br>AND THE AGE OF SEMICONDUCTOR TECHNOLOGY DOMAINS.....  | 71 |
| FIGURE 23: EMPIRICAL CUMULATIVE DISTRIBUTION OF THE PERCENTAGE OF YOUNG, NEW OLD AND<br>PERSISTENT OLD PATENTS FOR ALL THE DOMAINS IN THE PERIODS 1976-1985, 1976-1990,<br>1976-1995, 1976-2000 AND 1976-2006 ..... | 74 |
| FIGURE 24: THE LIFE-CYCLE SPACE .....   | 76 |
| FIGURE 25: ESTIMATED CUMULATIVE DISTRIBUTION FUNCTIONS FOR NEW AND INCUMBENT INNOVATORS ....  | 80 |
| FIGURE 26: MICRO-FOUNDED SPECIALIZATION INDEX FOR NEW AND INCUMBENT INNOVATORS.....   | 81 |
| FIGURE 27: ESTIMATED CUMULATIVE DISTRIBUTION FUNCTIONS FOR NEW INNOVATORS<br>FROM THE US, JAPAN, AND THE THREE ASIAN TIGERS .....   | 82 |
| FIGURE 28: MICRO-FOUNDED SPECIALIZATION INDEX FOR INCUMBENTS AND NEW INNOVATORS BY<br>GEOGRAPHIC ORIGIN .....   | 83 |
| FIGURE 29: THE EVOLUTION OF THE MICRO-FOUNDED SPECIALIZATION INDEX OVER TIME .....  | 84 |

|  |     |
|--|-----|
| FIGURE 30: THEORETICAL FRAMEWORK.....  | 92  |
| FIGURE 31: THE RELATIONSHIP BETWEEN INVENTORS AND INNOVATORS .....   | 97  |
| FIGURE 32: FICTITIOUS CITATION NETWORK WITH SHARES OF GENETIC HERITAGE .....   | 105 |
| FIGURE 33: DISTRIBUTION OF KNOWLEDGE BREADTH, DEPTH, PATENT COUNT AND EXPERIENCE .....   | 113 |
| FIGURE 34: SCATTER PLOT OF CONTINUOUS VARIABLES .....  | 115 |
| FIGURE 35: SCATTER PLOT OF CATEGORICAL VARIABLES VERSUS TECHNICAL KNOWLEDGE BASE SIZE .....  | 116 |
| FIGURE 36: ADJUSTED PREDICTIONS AND AVERAGE MARGINAL EFFECTS OF KNOWLEDGE BREADTH.....   | 124 |
| FIGURE 37: ADJUSTED PREDICTIONS OF RECENT KNOWLEDGE PERSISTENCE (MAIN TRAJECTORIES) .....  | 126 |
| FIGURE 38: EMPIRICAL CUMULATIVE DISTRIBUTION FUNCTIONS OF THE PROBABILITY OF CITING<br>A PATENT AS A FUNCTION OF ITS AGE BY CITING COUNTRY. .... | 151 |
| FIGURE 39: MAIN COMPONENT OF THE NMPs 1976-1980.....   | 153 |
| FIGURE 40: MAIN COMPONENT OF THE NMPs 1976-1985.....   | 153 |
| FIGURE 41: MAIN COMPONENT OF THE NMPs 1976-1990.....   | 154 |
| FIGURE 42: MAIN COMPONENT OF THE NMPs 1976-1995.....   | 154 |
| FIGURE 43: MAIN COMPONENT OF THE NMPs 1976-2000.....   | 155 |
| FIGURE 44: MAIN COMPONENT OF THE NMPs 1976-2006.....   | 155 |
| FIGURE 45: STATISTICAL RELATIONSHIP BETWEEN PFI, POP, PATHC INDICES AND FIRMS' PATENT COUNT. ....  | 156 |
| FIGURE 46: ARCHETYPAL LIFE-CYCLE OF A GIVEN TECHNOLOGY DOMAIN WITH RESISTANCE TO DECLINE.....  | 157 |
| FIGURE 47: ARCHETYPAL LIFE-CYCLE OF A GIVEN TECHNOLOGY DOMAIN STARTING WITH A DISRUPTION.....  | 157 |
| FIGURE 48: THE SPACE OF TECHNOLOGY DOMAINS BETWEEN 1976 AND 1980 .....   | 158 |
| FIGURE 49: THE SPACE OF TECHNOLOGY DOMAINS BETWEEN 1976 AND 1985 .....   | 158 |
| FIGURE 50: THE SPACE OF TECHNOLOGY DOMAINS BETWEEN 1976 AND 1990 .....   | 159 |
| FIGURE 51: THE SPACE OF TECHNOLOGY DOMAINS BETWEEN 1976 AND 1995 .....   | 159 |
| FIGURE 52: THE SPACE OF TECHNOLOGY DOMAINS BETWEEN 1976 AND 2000 .....   | 160 |
| FIGURE 53: THE SPACE OF TECHNOLOGY DOMAINS BETWEEN 1976 AND 2006 .....   | 160 |
| FIGURE 54: AGE STRUCTURE OF TECHNOLOGY DOMAINS .....   | 162 |
| FIGURE 55: EMPIRICAL CONDITIONAL PROBABILITIES OF SURVIVAL.....  | 174 |
| FIGURE 56: EMPIRICAL LEFT-CUMULATIVE DISTRIBUTION FUNCTIONS OF KB AND KD FOR SURVIVING AND<br>EXITING CORE INNOVATORS.....                       | 176 |
| FIGURE 57: CORE INNOVATORS' LOCATION IN THE KNOWLEDGE BREADTH AND DEPTH SPACE BETWEEN<br>1986 AND 1990 .....                                     | 184 |
| FIGURE 58: CORE INNOVATORS' LOCATION IN THE KNOWLEDGE BREADTH AND DEPTH SPACE BETWEEN<br>1991 AND 1995 .....                                     | 184 |
| FIGURE 59: CORE INNOVATORS' LOCATION IN THE KNOWLEDGE BREADTH AND DEPTH SPACE BETWEEN<br>1996 AND 2000 .....                                     | 185 |
| FIGURE 60: CORE INNOVATORS' LOCATION IN THE KNOWLEDGE BREADTH AND DEPTH SPACE BETWEEN<br>2001 AND 2006 .....                                     | 185 |
| FIGURE 61: COUNTRIES' KNOWLEDGE DIVERSIFICATION PATHS.....   | 186 |

## ACKNOWLEDGEMENTS

Doing a PhD is an enormous privilege. Studying for a PhD gives you the freedom of embarking in a journey of learning and discovery. Discovering what you want to know, learning how much you do not know and realizing how much you will eventually end up knowing. Perhaps more than any other “real” job (as it is common to hear), doing a PhD creates a psychological feeling which mirrors that of the least experienced of ancient times’ sailors embarking on an exploratory expedition to the unknown seas; the feeling of uncertainty ahead at the petering out of the sight of land, into the endless sea and dark cloudy skies. Scary but equally exciting. Then there is a deep feeling of pride at the end of the journey; the appreciation of how far you have come and the excitement of touching on land again. A land that you might have misjudged in size and location. A land whose existence may have been known before by others. A land that may even have been populated for quite a while. Regardless of the actual magnitude of your finding, such land is marvelous and the journey is unforgettable. So unforgettable that you now want to keep traveling and exploring. You realize that it does not really matter if you were the first one to achieve it, and neither does it matter if your destination is what you had planned for at the start. You learned to love the feeling of uncertainty about the destination and the route. You learned that looking for the right path might even be more rewarding than finding it. All of this is what a PhD thought me. If I feel that this journey was a privilege, it is certainly thank to the people that accompanied me throughout it and those that I met on the way.

First, and foremost, the Captain and the Co-Captain of the boat, who were brave enough to leave the helm in the hands of the young inexperienced sailor. My deepest gratitude goes to my supervisor, Bart Verspagen and my co-supervisor, Önder Nomaler, for providing comments and suggestions when I needed advice, for trusting me when I needed to build confidence, for being patient when I needed time and for being critical or sceptical when I needed to strengthen my findings. My thesis has greatly benefited from their supervision. As it has done from the prompt suggestions for improvement provided by the members of the assessment committee, the chairman Robin Cowan, Koen Frenken, Franco Malerba and Lili Wang, who took their valuable time to read the thesis under time pressure. For this and their insightful comments, I am very grateful to them.

Some of the chapters of this dissertation have greatly benefited from comments by Lap Chan, Roberto Fontana, Jeff Funk, Bronwyn Hall, Dominik Hartmann, François Lafond, Jianxi Luo, Chris Magee, Arianna Martinelli, Pierre Mohnen, Daniel Opolot and Jennifer Taborda. A special thanks to them for taking their invaluable time to discuss preliminary (and sometimes confusing) versions of this work.

Sometimes you learn something from people that may not even be aware of their teaching service. I would like to thank a number of professors with whom I regularly or occasionally interacted during my undergraduate and graduate studies and as a young researcher. Their scientific rigour and passion for the pursuit of knowledge taught me the high honour and societal value associated with the profession of academic research. These are Stefano Breschi, Nicoletta Corrocher, Giovanni Dosi, Francesco Lissoni, Patrick Llerena, Franco Malerba, Raffaella Piccarreta, Andreas Pyka, Pier Paolo Saviotti, Jerry Silverberg, Luc Soete, Shyama Ramani and Eddy Szirmai.

The start of this research journey was made possible by the financial support from the *Explore* project of the Dynamics of Institution and Markets in Europe (DIME) network of excellence, under the 6<sup>th</sup> Framework Programme of the European Union. Subsequently, I received financial assistance by the Maastricht School of Business and Economics. To both I am grateful for.



A very special mention goes to my amazing batch, Alejandro L., Daniel O., François, Samyukta, Jocelyn and Tatevik, with the invaluable and irreplaceable addition of Francesca and Jennifer. Thank you all for sharing almost every day of the entire journey and for being such an extraordinary group of friends, which made me feel at home in the ocean. You have been the best companions one could imagine for life- and research-related discussions; dinners and parties; frivolous and serious moments. This is a truly unusual combination of qualities. I feel extremely lucky that I met you and that we went through this journey together.

Without any doubt, there is one person that starts making you feel welcomed in Maastricht, even before you get there and keep doing so well after you left. A huge thank to Eveline for serving as secretary of the PhD program far beyond her professional duties, for taking such a good care of the PhD fellows and for always being helpful, during each stage of the PhD process.

PhD research and teaching life is definitely easier when you can count on an excellent administrative and technical support staff like the one at UNU-MERIT and the Maastricht School of Business and Economics. For that, I am grateful to Ad, Eric, Herman, Howard, Marc, Mourik Jan, Silvana, Sylvia and Wilma.

In addition, thanks to all other friends that I have met and interacted with along the way and that have significantly contributed to make my PhD a memorable journey: Ajay, Alejandro G., Alison, Andrea, Antonio, Augustin, Benjamin, Can, Carlos, Charlotte, Cheng, Claudio, Daniel V., Dominik, Edu, Elisa, Francisco, Hibret, Ilire, Iman, Jinjin, Jojo, Juan Carlos, Julietta, Lilia, Luis, Michael, Michiko, Omar, Paula, Sayan, Sergio, Shuan, Simone, Stefania F. and Stefania I.

When even your network of friends and colleagues cannot help, you usually revert to reading PhD comics and searching stackexchange.com, Matlab Central, Statalist.org, Wikipedia or looking at online open lectures. I am profoundly grateful to everyone who contributed to the growth of the internet as repository of collective knowledge.

The transition between being a PhD student and being a postdoctoral researcher can be tricky sometimes, especially when you happen to be both for a while. This process was enormously facilitated by my postdoc advisors at SUTD, Jianxi Luo, and MIT, Chris Magee, and by my colleagues and friends Bowen, Jeff, Riccardo and Ulf at the SUTD-MIT International Design Centre. Thanks for making me feel like a peer, pushing me to keep learning and for just have a lunch, coffee or dinner together to break the work routine.

Finally, and most importantly, there is no possible way to express the gratitude that I owe to my beloved ones, to which this thesis is dedicated. *Alla mia famiglia va il mio più grande ringraziamento per l'amore e il supporto che mi avete sempre fornito fin dalla nascita. Grazie per essermi sempre vicini anche se siamo lontani. Siete il baricentro che mi permette di rimanere in equilibrio anche quando a volte la testa inizia a girare. Spero che questa tesi possa in qualche modo ripagarvi dei vostri sacrifici e della prolungata distanza.* To my love, Jennifer, more than anyone else you made this voyage both possible and enjoyable. You make me confident when I am not. You can wipe off any fear with a smile or joke. You are always there, no matter if it is to share happiness or to listen to my grumpy complaints. You patiently listened when I was constantly talking about my thesis and always took time to discuss my findings. You had a look at a “countless” number of figures and always gave me the right suggestion or showed me an interesting point of view. You pushed me to believe in what I was doing. With you, every day is the happiest day. This allows putting things into their right place. Without this ability, completing my PhD would have been a far less joyful experience, if at all possible. *Gracias desde el profundo de mi corazón.*

## SUMMARY

The evolution of technology is a key driver of economic and societal change. To achieve a sound understanding of the process of economic development it is crucial to show how firms invent, learn and improve technology, and how technology and firms' strategies co-evolve. This thesis analyses firms' inventive strategies and technological catching-up in the semiconductor industry between the 1980s and the first half of the 2000s. Technological catching-up is intended here as the ability of new entrant firms to quickly reduce or revert the gap with the leaders and engage in research at the technological frontier, i.e. of affecting the direction of technological change. The Semiconductor industry is chosen as field of study because of the peculiar transformation of its industrial and technological ecosystem over the years. The industry was globally dominated by US firms until the end of the 1970s. It then witnessed the successful technological and market catch-up achieved by Japanese companies in the 1980s and, subsequently, by firms from South Korea, Taiwan and Singapore in the 1990s. In this thesis, we identify the strategies that incumbents and catching-up firms have followed when navigating through the technology space by means of statistical analysis of complex patent citations networks. We use patent data as proxies of inventions and patent citations as indicator of similarity in engineering challenges and problem-solving approaches.

This doctoral dissertation provides a new way of looking at catching-up, which departs from the usual focus on technological trajectories defined at the product level. We argue that products are not necessarily the correct unit of analysis for a study on firms' inventive strategies. The inventive process is largely an engineering problem-solving activity. Learning occurs by finding the right approach to tackle engineering challenges and overcome technical bottlenecks. Consequently, we choose to study paths of solutions (i.e. *engineering design trajectories*) on the space of engineering problems, which, potentially, can affect several products at once.

In Chapter 2, we provide a data-driven description of industrial dynamics and catching-up in the Semiconductor industry and highlight similarities and differences between successful and unsuccessful catching-up cases. Our findings show how the local Semiconductor industries in S.Korea, Taiwan, Singapore, Hong Kong, China and Malaysia shared similar initial national focus on semiconductors and strength of technical capabilities in this industry, up to the beginning of the 1990s. Yet, prominent catching-up countries (S.Korea, Taiwan and Singapore) later followed very different inventive strategies than less successful ones, especially regarding which technologies they improved upon.

In Chapter 3, we analyse in more detail the peculiar inventive strategies that have been followed by Korean, Taiwanese and Singaporean firms and how their actions affected the direction of technological change in the industry. We show that, up to the 1990s, firms from these countries prevalently devoted inventive effort to solving common engineering challenges by applying established approaches to problem solving. In other words, they followed the prevailing *engineering trajectories*. However, from the beginning of the 2000s, some of these firms began allocating more inventive effort to new engineering challenges now faced by the Semiconductor industry. This is especially true for those related to LCD monitor technologies and advances in metal-oxide-semiconductor field-effect transistors (MOSFETs).

In Chapter 4, we assess the evolution of the revealed technological advantage of incumbents and new entrants. To accomplish this, we develop a method to identify *technology domains* and their life cycle stage. Technology domains are defined as areas of research that share a set of common technological problems that are tackled by applying similar approaches. Our method is able to infer the life cycle stage of technology domains by analysing their engineering trajectory.

We then investigate in which domains new entrants and incumbents have a revealed comparative advantage. We show how new entrants from latecomer countries were able to upgrade their comparative advantage from relatively mature and declining technology domains to emerging ones in the period between the mid-1990s and mid-2000s.

In Chapter 5, we map firms' technical knowledge genome along two dimensions, depth and breadth. We do this by building on an analogy between knowledge and genetic evolution. We look at how much of the knowledge of technical solutions that has been generated historically in a given technology domain is inherited by a firm (i.e. the firm's knowledge depth), and how a firm's knowledge is distributed across domains (i.e. the firms' knowledge breadth). We hypothesize that firms' knowledge depth and breadth crucially affect their probability to persist innovating at the frontier. We show how the Semiconductor industry is characterized by a strong level of technical knowledge modularity. This means that what has been learned in a given domain is not particularly useful in other domains. Our findings show that, in this context, knowledge diversification is necessary to survive technological shocks that take the form of a change in the way semiconductor-related engineering challenges are tackled (i.e. a change in engineering trajectory). Firms that have a large knowledge breadth are more likely to persist innovating at the frontier after a change in engineering trajectory. This is because they have better chances that at least some of their current knowledge will still be useful in the future. This effect is positively enhanced by also having a large knowledge depth.

The theoretical framework and quantitative methods developed in this thesis open a path to better understand the fundamental cognitive drivers of technical change and to assess multi-dimensional technological determinants of firms' success. Knowing which firms are special and in which respect eventually makes it possible to investigate what makes them so. In the last part of the dissertation, we present a research agenda for unleashing the potential of this new theoretical and methodological approach for improving our understanding of the complex co-evolutionary pattern of technical change and industrial dynamics.



# 1 INTRODUCTION: COMPLEXITY, TECHNOLOGICAL CHANGE AND CATCHING UP

*"More than anything else technology creates our world. It creates our wealth, our economy, our very way of being"*

W. Brian Arthur (2009)

*"...agents, organizations and technology do not steadily proceed toward a global optimum. Rather they are linked to one another and their interactions create (and also prohibit) pathways for their future development"*

Magda Fontana (2014)

*"While innovation has been the main source of economic progress in the West, learning has also been important for the catch-up of the rest of the non-Western latecomers"*

Keun Lee (2015)

This PhD thesis provides a micro-founded and data-driven narrative of the successful catching-up in technical capabilities and market shares achieved by firms from latecomer countries in the Semiconductor industry. It does so by combining in an original way theories, concepts and statistical methods from technology dynamics, complexity, innovation, catching-up and network sciences.

Economic development can be described as a process of structural transformation and technological change characterized by discontinuities. There is a wide agreement in the literature on catching-up that some of the key drivers of growth rate differentials are technology and knowledge gaps and the presence (or the absence) of particular social and economic conditions that allow (or prevent) countries to absorb foreign technologies successfully (Fagerberg and Godinho, 2002, Fagerberg, 1994 and 1987, Abramovitz, 1994, Verspagen, 1991, Gerschenkron, 1962). Yet, catching-up has been mostly studied at the macro-level, despite being based on innovation and learning processes, which happen at the micro-level. Plenty of work provided anecdotal evidence on learning strategies applied by successful catching-up firms (e.g. Lee and Lim, 2001; Mathews and Cho, 1999, Cho et al., 1998) and related policies implemented by latecomer countries (e.g. Bell and Juma, 2008; Chang and Tsai, 2002; Chang et al., 1994). However, there is a scarcity of data-driven and firm-level studies that simultaneously analyse technology dynamics and firms inventive strategies, in the context of catching-up (with the notable exception of Lee (2014)). This PhD thesis contributes to fill this gap. We chose to study the Semiconductor industry because of the peculiar evolution of its technological and industrial ecosystem. It is a dynamic high-tech industry in which catching-up prominently occurred (Langlois and Steinmueller, 1999). Consequently, it provides a very interesting example of co-evolution of technological change, inventive strategies and catching-up. We deliver a data-driven statistical answer to two key questions: (i) *How did the technological environment evolve in the Semiconductor industry at the time that firms from latecomer countries successfully catch-up with the leaders?* (ii) *Which inventive strategies have been followed by successful catching-up firms?* Answering the former question allows understanding if any idiosyncrasy at the technology level favoured the catching-up process and if the technological landscape of the industry significantly changed after the emergence of new players. At the same time, revealing which inventive strategies have been followed by latecomer firms allow understanding how they interacted with the technological landscape during the catching-up process. The reader might wonder, what do we mean by inventive strategies? Throughout this work, we intend them as a multidimensional concept. In its essence the term ‘inventive strategies’ refers to the common denominator that links together the following questions: *which technologies do you improve upon? Are they local or foreign-invented? How quickly do you go through improvement cycles? Do you venture to upgrade your skills to emerging technology domains or you rather strengthen your comparative advantage in mature technologies? Do you opt to specialize in a narrow set of technologies or do you attempt building a broad technical knowledge base?* In this thesis, we answer these questions by analysing patent data, which are the largest available records of technology developments and provide information on firms’ inventive histories. To do that we design novel analytic tools that extract significant information out of noisy, complex and dynamic citation networks.

Since, in this work, we focus our analysis on inventive strategies and catching-up, we need to ask ourselves an additional crucial question: *Are there universally valid and replicable catching-up recipes?* It is very tempting to search for generalizable lessons for rapid economic development. This exercise is, however, prone to oversimplifying the problem. This is due to the nature of the co-evolution of the technological landscape, the industrial ecosystem, the strategic decisions made by the actors that populate them and their capabilities. Long run market success

is ultimately based on the development of frontier technological capabilities (Lee and Lim, 2001). Technological capabilities are defined as the knowledge and skills that an agent needs to create, use, adapt and improve technology (Lall, 1993)<sup>1</sup>. Technological capabilities are necessary to adapt to changes in the technology environment. At the same time, their development is also a direct source of technological change. Consequently, any meaningful theory, explanation or discussion of industrial dynamics and catching-up needs to be tied to the discussion of how the technical system underlying the industry or products of interest evolves and how inventive agents interact with the technology landscape. Thinking in terms of complexity and evolution helps improving our understanding of these interconnected phenomena (Frenken, 2006). It is also crucial to understand whether general catching-up recipes may exist.

There are at least three key features of technological change that make it a complex phenomenon: *uncertainty*, *interactions* and *path-dependence*. Neglecting these features might lead to misleading conclusions about the existence of catching-up recipes. The process of technological innovation can be seen as searching for better technical solutions in a technology space (Silverberg and Verspagen, 2003, Fleming and Sorenson, 2001). As such, the outcome of technological progress, for instance which features of existing technologies are improved or which new technologies are developed, is characterized by uncertainty. Depending on the magnitude of technical change and the level of heterogeneity of players' inventive strategies, this outcome can also be unpredictable. In its mildest form, technical progress is a stochastic process in which players do not know the best inventive strategies to follow at a given point in time, but they are able to associate probabilities to the different future directions of improvements. In this case, if we could observe several realizations of the same system starting from the same initial conditions, we would converge to an expected direction of improvements, even though individual runs might look very different. This would mean that the same inventive trajectories and catching-up pathways would be replicable *in the same context*. However, the arrival of radical innovations that revolutionize a technology domain or create an entirely new one, is usually characterized by Knightian uncertainty. This makes it impossible to even define the probabilities of the different outcomes. To some extent, innovations of such a large scale are idiosyncratic strokes of individual genius sometimes with a strong component of luck, like the discovery of Penicillin by Alexander Fleming. Even when they are the results of a collective process, radical innovations need a particular combination of talented researchers and engineers, forward-looking managers and risk-taking investors that make their realization a process that is highly sensible to small perturbations and, therefore, chaotic in nature. In this case, even rewinding the system and running it again might lead to different outcomes. In a Knightian kind of world, replicable catching-up recipes cannot exist. However, radical innovations and paradigm shifts are rare events (Fleming, 2007, Silverberg and Verspagen, 2005 and 2007) and a lot of novelty is cumulated through incremental improvements whose trajectory can be subject to educated guesses. The most famous of them is perhaps Gordon Moore's law. Moore, one of Intel's founder, predicted in a 1965 paper that the number of transistors on a chip would have duplicated every 18 months (Moore, 1965), as it indeed happened since then.

A lot of the uncertainty intrinsic in the process of technological change comes from the interactions between players' inventive strategies, the technological structure of a given industry

---

<sup>1</sup> They are also known as technology-based design capabilities or engineering design capabilities (Luo et al., 2012; Dym et al., 2005).

and the technology and industrial policies that different governments overlay on the system (Malerba and Orsenigo, 1997). These interactions make the trajectory of technology improvements and catching-up pathways context specific. Similar strategies might lead to dissimilar outcomes in different national and sectoral contexts. Moreover, there are many different sorts of interdependencies between technologies which make them *complex adaptive systems* (Frenken, 2006; Foster, 2005), or, as defined in the engineering literature, *complex engineering systems*, which are rich in human and technology complexity (De Weck et al., 2011). Technologies have within, between and inter-temporal connections. Current technologies are related with their previous generations by a relationship of improvement and are connected to current generations by a relationship of possible cross-fertilization. Technical progress in a given technology can spill to others depending on the relatedness of the engineering problems driving its improvement. Arthur argues that “*new technologies are not ‘inventions’ that came from nowhere [ . . . ] They are created, - constructed, put together, assembled – from previously existing technologies*” (Arthur, 2009, p. 2, words in regular font style are adapted from the original quote).<sup>2</sup> Knowledge spillovers and cross-fertilization effect can be better understood if we consider technologies as part of a nested system, or, borrowing once again Brian Arthur’s words, as “*systems of technologies*” where “*each component of technology is itself in miniature a technology*” (Arthur, 2009, p. 23). The interdependence between different components reveals the structure that explains why technical advances differently diffuse across technologies and why technological opportunities are not evenly distributed across industries and domains, and are clustered in time. This ultimately affects catching-up prospects of latecomers and casts doubts on the existence of context independent catching-up recipes.

The nested nature of technologies and the direction taken by past improvements are ultimately responsible for the third key feature that make technological change a complex phenomenon, namely path-dependence. As argued by Foster (2005), “*such a system must exhibit some degree of structural irreversibility due to the inherent hierarchical and ‘bonding’ nature of the connections between components that are formed as structural development proceeds. It is this that results in the inflexibility and maladaptiveness that precipitates a structural discontinuity of some kind.*” (Foster, 2005, p.3) From the perspective of catch-up, the structural and intertemporal connections across technologies not only channel novelty along rigid pathways, but also make firms’ learning choices irreversible. Several authors stressed the relationship between the degree of cumulativeness of technological change and the ease of entrance of new players (Lee and Lim, 2001; Malerba and Orsenigo, 1997; Dosi, 1982). The important lesson is that *the irreversibility of the learning trajectory makes catching-up reversible*. What appears to be a lexical oxymoron is actually the reality that latecomer firms face. Learning along the wrong direction, perhaps due to initial positive market and technical feedbacks, can lead latecomer firms to get stuck and interrupt the process of technical capabilities upgrading. This can push countries into the middle-income trap or even make them regress along the income ladder (Lee, 2014). Furthermore, wrong learning paths can prevent innovation systems to complete the transition from a learning-based system to a novelty generating-one (Vertesy, 2011). If learning is path-dependent and economic development is based on knowledge-upgrading then the latter is also strongly path-

---

<sup>2</sup> Similar thoughts on the existence of knowledge spillovers and percolation effects, have been elaborated by, among others, Kauffman (1988), Verspagen (1997), Silverberg and Verspagen (2005; 2007a; 2007b), Krafft and Quatraro (2011) and Metcalfe (2014).



dependent. The importance of *looking for the right path* is therefore crucial for the growth prospects of countries and firms.

Path-dependence can be a blessing or a curse. Obviously, incumbents favour cumulative technical improvements as long as they follow the direction that rewards the knowledge that they accumulated. When that is the case, incumbents enjoy a large advantage over latecomers. The emergence of discontinuities brakes that advantage. Thus, the faster those discontinuities arise (i.e. the shorter the technology life cycle), the more windows of opportunities for new entrants open up (Lee, 2014). Path-dependence also plays against the effectiveness of applying what has been learned in one national and sectoral context to different ones. Systems that are subject to path-dependence have an intrinsic resistance to radical changes. For this reasons, applying a set of policies or following a specific inventive strategy just because it has been successfully applied elsewhere has no guarantee to succeed, unless the given industry or firms share similar histories with those intended to be imitated.

The characterization of technology and catching-up dynamics as complex adaptive systems that we sketched here raises scepticism on the existence of universally valid and replicable catching-up recipes. Does this mean that we are condemned to ignorance? Does this imply the irrelevance of studying these phenomena? Absolutely not. Characterizing the dynamics of economic and technical systems as complex phenomena must not be mistaken as stating that they are purely random processes (Baofu, 2007). In contrast, thinking in terms of complexity helps analysing the emergence of patterns and study their properties. The essence of complexity thinking applied to technology and catching-up dynamics lays in understanding two crucial messages. First, one should devote great efforts in separating signal from noise and identify statistically significant peculiarities in terms of inventive strategies followed by successful players along their development path. Second, this exercise should attempt to unfold history by revealing what makes successful latecomers *really* special, rather than trying to search for recipes in which to trust blindly. This allows revealing historical micro-level trajectories of catching-up. These two messages guide the analysis developed in this thesis.

## 1.1 Conceptual framework and structure of the thesis

The thesis analyses technological catching-up in the semiconductor industry at the firm level. Semiconductors are the source of catching-up narratives *par excellence*. As will be discussed in Chapter 2, this industry has seen a profound change in its industrial ecosystem, with the successful technological and market catch-up of firms from latecomer countries like South Korea, Taiwan and Singapore. In this thesis, by technological catching-up we mean the ability of new entrant firms to engage in research at the technological frontier, i.e. of affecting the direction of technological change.

The three main characteristics of technology that makes it a complex adaptive system explained above, namely *uncertainty*, *interdependence* and *path-dependence*, have pushed scholars to focus mainly on the *direction*, *cumulativeness* and *speed* of technical change as drivers of catching-up processes. These dimensions have proven to be key determinants of the ease of entry (Lee and Lim, 2001, Malerba and Orsenigo, 1997, Klepper 1997 and 1996, Dosi, 1982). However, we argue that, if we conceptualize technologies as part of a structured system in which opportunities of advances are unevenly distributed and irregularly connected, as done by Arthur (2009), Silverberg and Verspagen (2005) and Thurner et al. (2010), four additional properties of the system of technologies become crucial in the context of catching-up: the *structure* and *ranking* of the system of technologies, the magnitude of change in problem-solving approaches

(which we later define as engineering trajectories) and the *heterogeneity of life cycle stages* of technology domains. Throughout the thesis we intend technology domains as areas of research that define a set of common technological problems that are tackled applying similar mindsets and toolboxes. A high-tech industry like the Semiconductors is characterized by an heterogeneous set of technologies. Technologies are solutions to particular needs. Satisfying those needs requires solving given technical challenges. Engineering problems can be tackled from a variety of approaches. As will be extensively discussed in Chapter 3, 4 and 5, the sequences of improvements within a given technological domain highlight which approaches to problem solving have been used. We define these technical paths as *engineering design trajectories*, which are the projection in the engineering design space of trade-offs between technical characteristics determined at the product level, i.e. of *technological trajectories* (Dosi, 1982). Decisions on which features of a given product need to be improved and how the product is intended to be designed define which engineering challenges need to be overcome. Throughout this thesis, we will also discuss how the similarity in the kind of technical problems faced in an industry generates the structure of the system of engineering problems. Engineering challenges can be interconnected and solving bottlenecks opens up large and widespread opportunities for technical improvements. The structure of the system determines the availability of technological diversification pathways and, ultimately, the likelihood of being a persistent core innovator (Chapter 5). The heterogeneity of beliefs and search strategies in the technology design space, at the firm level, determines the simultaneous existence of different approaches to problem solving and the level of inventive efforts devoted to different problems (Chapter 3). Therefore, engineering problems can be ranked based on how much inventive efforts they attract. The level of inventive effort is also related to the breadth of problem-solving approaches applied to tackle engineering challenges. Problems that are collectively determined to have high priority attract a lot of inventive effort that is likely to cover a wide spectrum of approaches. Therefore, emerging technology domains are likely to be characterized by heterogeneous beliefs on which are the most promising approaches to solve a technical problem (Chapter 4). Eventually market selection will foster convergence to a subset of search strategies. However, the entrance of new players may bring back a large variety of approaches to problem solving into the system. Therefore, as will be shown in Chapter 4, technology domains have life cycles, and, at each point in time, there will be several domains at different stages of their life cycles within an industry.

The structure and ranking of the system of technologies, the magnitude of change in approaches to problem solving and the heterogeneity of life cycle stages, play an important role in determining the ease of catching-up. This role is related to the different ways in which they affect path-dependence. Suppose that the structure of the system of technologies is such that different technology domains and their components are largely and equally connected to each other in terms of knowledge relatedness. Suppose also that change is localized in one or few components. In this case, the probability of breaking path-dependence in the learning advantage of incumbents is at its lowest level. The capabilities that incumbents have developed in the past are still useful for their survival, or at least allow them to diversify into areas not affected by change in the approaches to problem solving. When change is still localized but the system is more sparsely connected then the probability that a small window of opportunity for latecomer opens up is slightly larger. This is because now incumbents specialized in the components subject to change have less chances to use their accumulated knowledge in related areas. Suppose now that the magnitude of changes in the approaches to problem solving is larger and pervades to several domains. If the system is highly connected, incumbents might still enjoy a small advantage but only provided that are able to successfully recombine their knowledge. If the system

is sparsely connected then path-dependence is fully broken and a pathway of entry opportunities lay before latecomers. Similarly, the existence of multiple technology domains at different stages of their life cycle also interacts with firms' inventive histories and their strategic options for the future. New entrants can decide to focus on emerging domains and enjoy the potential early-mover advantage (but also face the hazard) associated with being the first to tackle new engineering challenges or explore new problem solving approaches. They can also decide to focus on more mature or declining domains, in which niches with less competition might be available for entrance and provide a safe environment for technical learning.

The novel theoretical contribution of this thesis resides in making the first attempt toward including the structure and ranking of the system of engineering design problems, the magnitude of change in approaches to problem-solving and the heterogeneity of life cycle stages of technology domains in the discussion of technical catching-up. We use a combination of methodologies based on patent citation network analysis to study catching-up and inventive strategies in the semiconductor industry. The thesis is composed of four essays that are theoretically and methodologically intertwined. We analyse patent citation networks to identify the main engineering design trajectories followed by the industry. This is done by applying a methodology that identifies the network of main paths (NMPs) within large citation networks (Vespagen, 2007, Martinelli, 2010, 2012; Hummon and Doreian, 1989). More specifically, citations are interpreted as footprints of similarity in approaches to problem solving. The NMPs highlights the sequences of citations that are more central across the *system of technologies*. The main path approach is then complemented by two other methodologies. We apply network community-detection techniques to identifying sub-technology domains in the industry, their structural relationship and the stage of their life cycle. Finally, we make use of a genetic approach to patent citation networks (Martinelli and Nomaler, 2014) to measure persistence and modularity of technical knowledge and the breadth and depth of firm's technical capabilities. In this approach, the knowledge of a specific solution to a selected engineering problem, which is represented by a given patent, is seen as a gene that spreads to the population of inventors. Citations therefore represent genetic heritage that persists from parent (the cited patent) to offspring (the citing patent). As a note for the reader, we stress that the methodologies developed in this thesis are also nested. The main path approach is used in Chapters 3, 4 and 5 as the basic building block, whose function is to reduce large citation networks to their backbone. Then further analysis is done *on top of* the network of main paths. Yet, despite the risk of redundancies, we decided to briefly introduce the whole procedure from the very beginning each time such that each chapter can be read as a stand-alone essay.

In Chapter 2, we provide a data-driven description of industrial dynamics and catching-up in the Semiconductor industry and highlight similarity and differences between successful and unsuccessful catching-up cases. In particular, we reveal how the Semiconductor industries in S.Korea, Taiwan, Singapore, Hong Kong, China and Malaysia shared similar initial national conditions at the technology level, up to the beginning of the 1990s. However, we unfold how they later followed very different inventive strategies in terms of source and speed of technical improvements. In the second essay (Chapter 3) we investigate whether we observe stability or discontinuity in the approach to problem solving and in the ranking of engineering problems of semiconductor technologies between 1976 and 2006. We also analyse which firms were pushing the trajectories toward either stability or discontinuity and how latecomer firms behaved in this respect. Findings show that the innovative entrance of firms from Korea, Taiwan and Singapore occurred in the mid-1990s in a moment in which the technological trajectory was stable. Up to the end of the 1990s, their learning strategy was prevalently focused on following long-

established paths of technical improvements and allocating their inventive efforts in accordance with the then prevailing ranking of engineering problems. Yet, in the beginning of the 2000s, some of them became active ranking-changers and start focusing on a new set of engineering challenges related to LCD monitor technologies and advances in metal-oxide-semiconductor field-effect transistors (MOSFETs).

Although very successful in terms of quickly acquiring influence on the direction of the engineering trajectories, the path following type of entrance applied by latecomers, raises a fundamental question concerning the sustainability of their specialization patterns. *Did latecomers succeed in upgrading their comparative technical advantage from relatively old technology domains to young and emerging ones?* This question is answered in the third essay (Chapter 4), where we develop a methodology to identify sub-domains of semiconductor technologies and categorize the stage of their life cycle. Our results show that, up to the end of the 1990s firms from Taiwan, Korea and Singapore specialized mainly in mature and exhausting technologies, sometimes attempting to renew them, whereas American and Japanese firms were comparatively better in younger areas. This learning process allowed capability building by Korean and Taiwanese latecomers that resulted in a revealed technological advantage in emerging areas at the beginning of the 2000s.

These results lead to a series of further research questions that are analysed in the fourth and last essay (Chapter 5). The fragmentation of the semiconductors' value chain from the 1990s onwards increased the degree of functional specialization of the players involved. Nowadays design, manufacturing, testing and commercialization of semiconductor devices are made by different firms. However, costs, profit margins and the level of competition differ widely across different stages of the value chain. In particular, the manufacturing stage of the semiconductor industry shows increasing costs (due to increasing product complexity), decreasing profit margins and increasing competition, the latter mainly due to the entrance of China. Therefore, to escape the low profitability and high competition trap Taiwanese, Korean and Singaporean firms need to move along the semiconductor value chain. The ease of this specialization upgrading crucially depends on their ability to diversify their technical knowledge. This is what we analyse in Chapter 5, where we study firms' innovative survival in the semiconductor industry as a function of knowledge diversification and accumulation. As we discussed earlier, the semiconductor industry is characterized by high levels of technical modularity. We hypothesize that, in such context, diversification and knowledge accumulation provide better chances for innovation persistence. We argue that, when knowledge is hardly portable across domains, the spread of survival chances between diversified and specialized players should be larger when technological change is knowledge replacing. Survival as core innovator (i.e. *innovative survival*) is defined as persisting having patents located on the main paths of citations. Our findings show that diversification significantly improves core innovators' odds of survival but only for large levels of knowledge breadth. The effect is indeed conditional to the level of knowledge persistence. In contrast, knowledge depth alone does not affect survival as core innovator. However, there are positive complementarities between knowledge breadth and depth.

# 2 TECHNOLOGICAL AND INDUSTRIAL EVOLUTION OF THE GLOBAL SEMICONDUCTOR INDUSTRY

## ABSTRACT

In a world pervaded by Information and Communication Technologies (ICTs), the Semiconductor industry is an engine of economic growth. It is also one of the few high-tech industries that witnessed a significant global diffusion of inventive activities. Several Asian countries attempted to narrow their technological gap with the leaders in this industry. Only three of them succeeded. These are S.Korea, Taiwan and, albeit to a smaller extent, Singapore. In this chapter we provide an overview of how the technological and industrial evolution of semiconductors. We quantify the technological and market success of these countries and discuss how their entry was favoured by the increase in the technical modularity of semiconductor devices. We also investigate similarities and differences of initial conditions between successful and unsuccessful catching-up countries. Finally, we show how the former markedly differ from the latter in the extent to which, after focusing on improving foreign technologies, they improved upon national ones. Firms from successful catching up countries also went through significantly faster cycles of technical improvements than their competitor in the US, Japan and the unsuccessful latecomers.

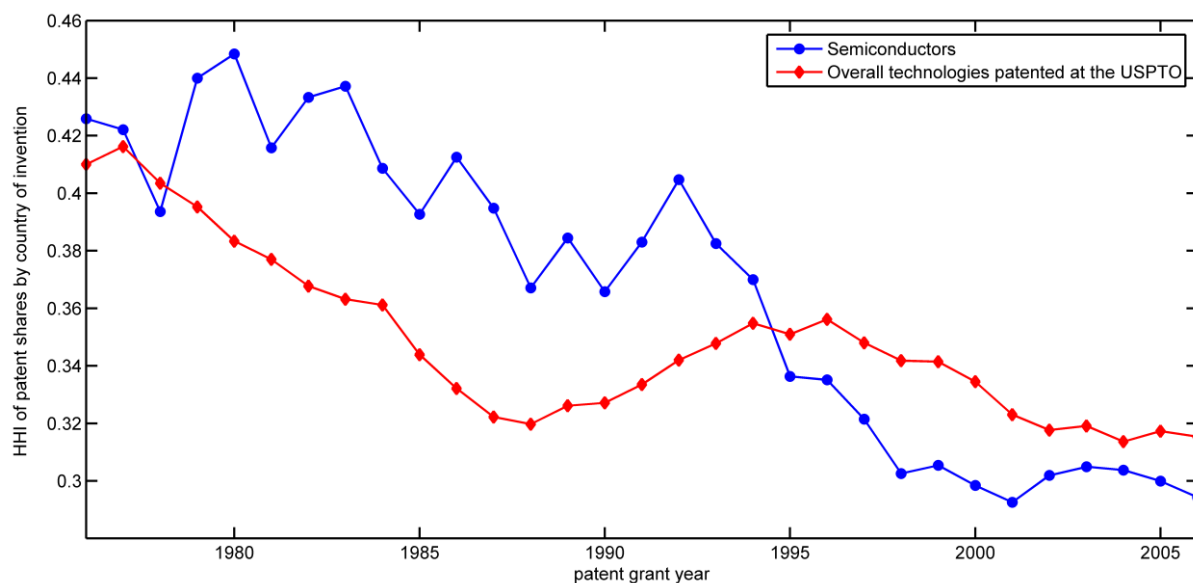
## 2.1 Introduction

Technology shapes the economy. The success and failure of firms, industries, and countries depends on their ability to keep pace with technological change and being active agents of it. To do so they need to constantly introduce, learn and adopt better technologies. Some technologies have wider effects on the economy as a whole. We currently live in a world shaped by information and communication technologies (ICTs), whose broad applicability has pervasive effects on the economy (Jovanovic and Rousseau, 2005, Soete, 2000 and Freeman and Soete, 1997). The flagship of ICT technologies is represented by the Semiconductor industry, whose products are ubiquitous in modern electronics. Improving semiconductor technologies is not only a source of pervasive technical change but also a key engine of economic growth. For instance, the U.S. semiconductor industry's share of U.S. GDP is the third largest of all U.S. manufacturing industries, behind only the petroleum refinery and pharmaceutical preparations industries (Parpala, 2014). Interestingly, the Semiconductor industry is one of the few high-tech sectors that spread to countries that were still experiencing a large technological gap with the leading ones, at the time of their entrance in this industry. Given semiconductors' pervasive role as general-purpose technologies and its peculiar globalized industrial structure, it is natural to wonder how such inclusive growth has occurred. Did the evolution of semiconductor technology favoured the entrance of new players? What does it make successful catching-up firms special?

In this chapter, we provide an overview of the evolution of the technological and market evolution of the Semiconductor industry. We show how this industry has been characterized by inclusive growth of technological innovations since about the beginning of the 1980s. We also discuss how technical and market catching-up has prominently occurred in two waves. In the first wave, which happened in the 1980s, market and technological advantage of US firms has been eroded by the entrance of Japanese companies. In the 1990s, a second wave of entrance from Asian Tigers' firms disrupted market structure. The progressive development of strong technical capabilities allowed catching-up firms to rapidly gain market shares. We discuss how the evolution of the industry's product portfolio and underlying technologies, in particular the increase technical modularity, have favoured this catching-up waves. By analysing patent records, we also provide a preliminary analysis of the inventive strategies followed by firms from successful and unsuccessful catching-up countries in terms of the sources and speed of technical improvements. We also show how, in successful catching-up countries, the inventive activities of local firms have led to the formation of a national industry, with several players that progressively improved upon each other's technologies.

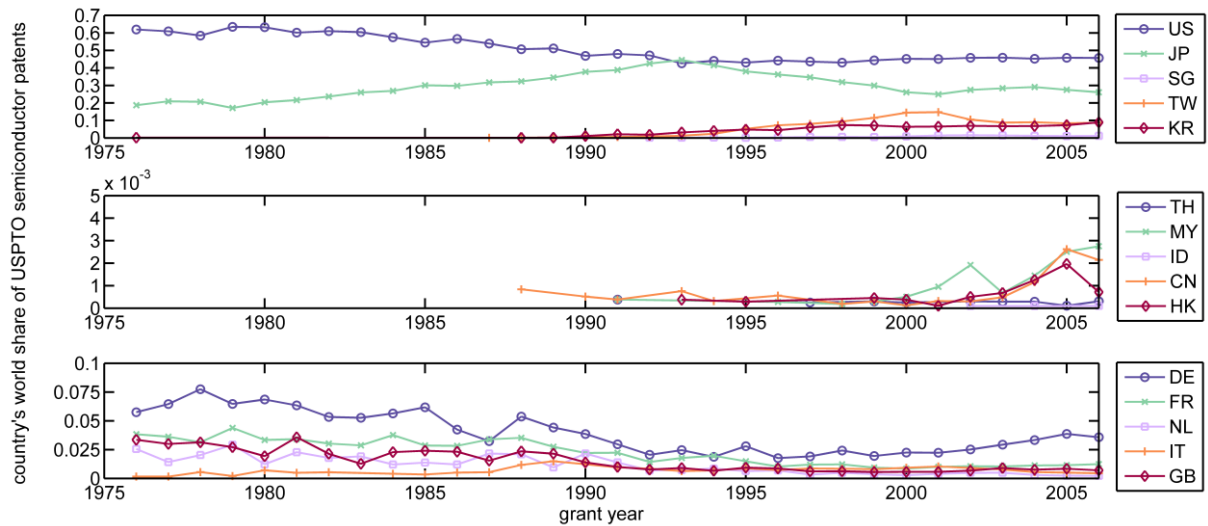
## 2.2 Technological change, industrial dynamics and catching-up

We study the semiconductor industry because it provides a unique case study to test theories and empirical evidence of the interplay between technological change and catch-up. Until the mid-1990s, Semiconductor inventive activities used to be more concentrated geographically than expected by looking at the overall set of technologies patented at the United States Patent and Trademark Office (USPTO), as measured by the Herfindahl-Hirschman concentration index (HHI) of semiconductor patent shares by country of invention reported in Figure 1. Then, Semiconductor inventive activities spread out globally considerably more than expected.



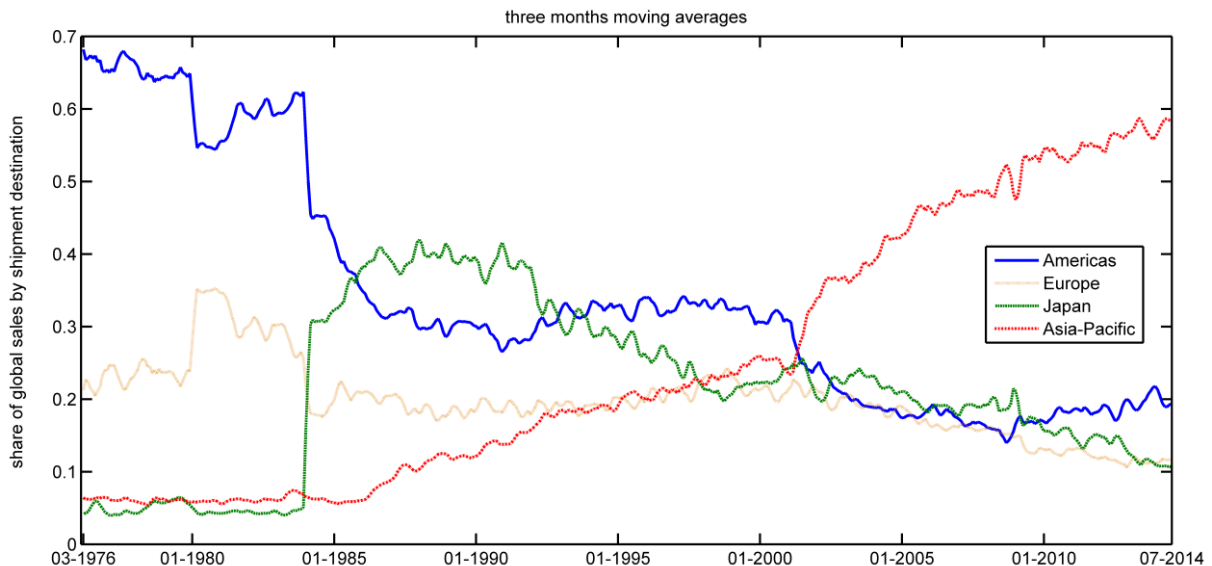
**Figure 1: The globalization of Semiconductor inventive activities**

It is natural to wonder then, which geographical locations have become important sources of semiconductor inventions over time. Figure 2 shows the evolution of semiconductor patent shares by country of invention. In the 1960s (not reported in Figure 2) US firms held about 85% of patents granted by the USPTO in the semiconductor technological classes. At the beginning of the 1990s, this share collapsed to less than 50%. This decrease was largely due to the rise of Japan as the US' main competitor. However, just after matching the US patent share in 1993, Japan's one starts decreasing, while the US level remained approximately constant. Japanese competitive edge as top semiconductor innovator was eroded by the catching-up in technical capabilities by S.Korea, Taiwan and, albeit to a smaller but still non-negligible extent, Singapore. Hobday (1995), discussed how a wider set of latecomer countries started manufacturing Semiconductors at the beginning of the 1990s. These are China, Hong Kong, Indonesia, Malaysia and Thailand. As a comparison, we report their patent shares. Their performance up to 2006 has been poor. China, Hong Kong and Malaysia, have recently experienced growth in their Semiconductor patent shares, while Indonesia and Thailand's shares have been languishing. However, despite this recent growth, patent shares for China, Hong Kong and Malaysia were still very marginal, until at least 2006, compared to the technological leaders (US and Japan) and the main followers (Korea and Taiwan). They were also considerably lower than any of the top European countries (Germany, France, Italy, Netherlands and UK), which are themselves very marginal contributors to technological innovation in the Semiconductors.



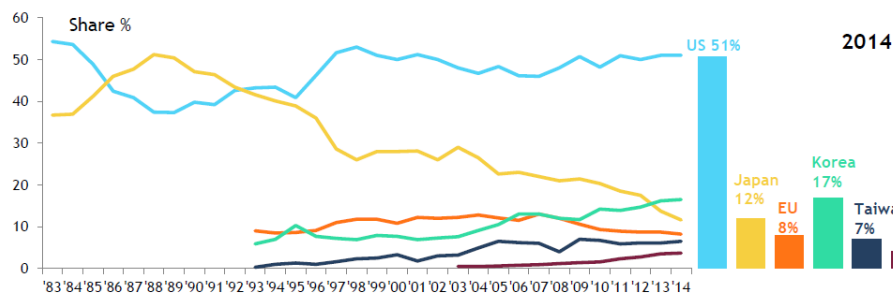
**Figure 2: USPTO semiconductor patent shares by country of invention**

Note: author's elaboration based on the NBER-USPTO data (Hall et al., 2001)



**Figure 3: Share of semiconductor sales by geographic destination**

Note: author's elaboration of data from the Global Sales Report (GSR) by the World Semiconductor Trade Statistics (WSTS). The GSR is a three-month moving average of semiconductor sales activity by billing country. The data is provided by the WSTS, which represents approximately 55 semiconductor companies worldwide.



**Figure 4: Share of semiconductor sales by geographic origin**

Note: Figure from Yinug (2015). Market shares based on headquarter of seller. Numbers rounded. Data source: SIA/World Semiconductor Trade Statistics (WSTS)/HIS/PricewaterhouseCoopers/IC Insights.



The catch-up in terms of technological output by Korea, Taiwan and Singapore, is closely matched by the rise in sale shares. Figure 3 shows how remarkable the similarity between the trends for patent shares by country of invention and the sale shares by billing country is. From the beginning of the 2000s the Asia Pacific region became the most important market in the world for semiconductor sales<sup>3</sup>. It is much more difficult to obtain comprehensive disaggregated data for semiconductor sales by country of origin. However, for those at our disposal<sup>4</sup>, which we show in Figure 4, we can notice a strong similarity between the trends for market and patent shares by country of origin<sup>5</sup>. The large correlation between these trends suggests that, as argued by Lee and Lim (2001), long run market success and improvements of technological capabilities co-evolve.

Semiconductors then clearly stand out of the crowd in terms of prominent entrance of new players from latecomer countries. This suggests that there may be something peculiar to the path of technological change in this industry that lowered barriers to entrance. The product portfolio of the industry could be broadly defined by two main categories: commodity integrated circuits and application specific integrated circuits (ASICs). The former includes all sort of integrated circuits (ICs) that are mass produced in a standardized manner and used in a large number of applications. Microprocessors, memories, simple standard chips, and transistors belong to this category. The second large category of devices is made by application specific integrated circuits (ASICs). The ubiquity of electronics in modern manufacturing process and products often required, besides standardized microchips, specific devices tailored to the computing needs of the given application. ASICs are therefore a very heterogeneous category that includes both ICs that are programmed by the producer to suit a specific need required by the user, and field-programmable gate arrays, which are designed to allow users to customize them to perform a specific function. Despite the undifferentiated nature of commodity devices, economic returns and the capabilities required to design them differ widely within the category. The microprocessors segment is largely dominated by Intel and Advanced Micro Devices and no player from latecomer countries managed to enter and consolidate in this segment. Memories make the opposite case. They are relatively easier to design and produce than microprocessor. Competition in this segment is therefore based on cost-advantage. As discussed by Langlois and Steinmueller (1999), this made it a particularly favourable area of entry for latecomers with initially limited technical capabilities but strong wage advantage. As we have shown in Figure 2, latecomers' entrance in the Semiconductor technological arena, clustered in two distinct waves. First Japanese producers successfully entered in the memory market and challenged the US leadership in the 1980s. Then large conglomerates from South Korea in particular, and Taiwan to a smaller extent, like Samsung, LG and Hyundai Electronics, prominently entered into the memory-making market.

ASICs are a more interesting case, which show the existence of crucial interactions between technological change and industrial dynamics even more prominently than memories. The diffusion of ASICs became possible thanks to the persistent miniaturization of integrated circuits components, as predicted by Moore's law, and the introduction of a particular design of ICs

---

<sup>3</sup> The data comes from the US Semiconductor Industry Association's Global Sales Report (GSR) based on data from the World Semiconductor Trade Statistics (WSTS) organization.

<sup>4</sup> The data source is Yinug (2015).

<sup>5</sup> Note that the SIA report excluded foundry output. Consequently, sales of Taiwanese and, partly, Chinese companies, are likely to be largely underestimated.

called complementary metal-oxide-semiconductor (CMOS). On the one hand, the miniaturization of chips made them ubiquitous and increased computational power. This also created demand for customized ICs. On the other hand, CMOS allowed high-density of logic functions on a chip, decreasing production costs. The two effects made it possible to invest in the design and manufacturing of ICs for small market segments that would have previously been unprofitable. This allowed the entrance of small specialized players which focused on specific technical skills and thorough knowledge of users' needs (Adams et al., 2013; Fontana and Malerba, 2010; Ernst, 2005; Linden and Somaya, 2003).

The combination of the emergence of CMOS technology, the diffusion of ASICs and the vertical disintegration of the value-chain, explain the second wave of globalization of the semiconductor industry, with the emergence of key players in the Asian Tigers (S.Korea, Taiwan, Singapore and Hong Kong), and assembly and test plants in Southeast Asia (Malaysia and Thailand) and China. Adams et al. (2013) explain how *"the increased adoption of Complementary Metal Oxide Semiconductor (CMOS) production processes weakened the interdependence of product design and manufacturing. [...] With the creation of standardized interfaces between components and Electronic Design Automation (EDA) tools a modular system developed. [...] The interdependence between product design and manufacturing was weakened in many product segments in semiconductors and specialist firms were able to enter the industry at both the design and the manufacturing stages"* (Adams et al., 2013, p.287). Technical modularity is a form of design that makes components of a system interdependent within modules but independent across them. (Baldwin and Clark, 2000). Modular unit can therefore be designed and manufactured in separate steps and locations. The emergence of modular designs is one of the forces that fragmented the value chain and fostered specialization in the semiconductors industry, in other words, technical modularity lead to organizational modularity (Ernst, 2005a, 2005b). However, modularity has not necessarily reduced complexity. As discussed by Ernst technical modularity allowed combining more and more layers of components into a single chip, making the design of the whole system more complex. In fact, since mid-1990s, ASICs have become so complex to include what were previously seen as undifferentiated devices like microprocessors and memory blocks into a single chip that performs the same functions of an entire system. These chips are indeed known as "System on a Chip" (SoC). The design, manufacturing and commercialization of complex SoC is certainly the most heterogeneous semiconductor market segment in terms of players, strategies and product characteristics (Linden and Somaya, 2003). As a way to reduce uncertainty due to the increased design complexity of SoCs, incumbent firms became quite conservative concerning exploring new design methodologies to avoid discovering problems in the subsequent manufacturing of large yields. This has open a room for risk-taking latecomers *"willing to use chip designs that they believed could help them to capture market shares"* (Ernst, 2005a, p.60).

The increased design, testing and manufacturing complexity is only one of the technical and business challenges faced by players in the semiconductor industry. Shorter product life cycles, shrinking profit margins, and global competition are additional key concerns (Brown and Linden, 2009). This is in particular true for the market segments and the steps of the value chain traditionally targeted by latecomer. Starting from the mid-2000s, the entrance of Chinese companies in the memory market and the testing, manufacturing and assembly of semiconductor devices has directly challenged the success of previous latecomers. Furthermore, while certainly providing entry opportunities for Asian Tiger's latecomers, as argued by Lee (2013), the shortening of the technology life cycle also threatens their ability to sustain their growth path, especially when facing the competition of more cost-effective Chinese firms. Therefore, sustaining technological advantage in the industry requires firms to be alert of the direction of technologi-

cal change, constant upgrade their technical knowledge to the latest design and manufacturing technologies and to be open to the use of new materials and production equipments<sup>6</sup>.

## **2.3 Are successful catching-up countries special?**

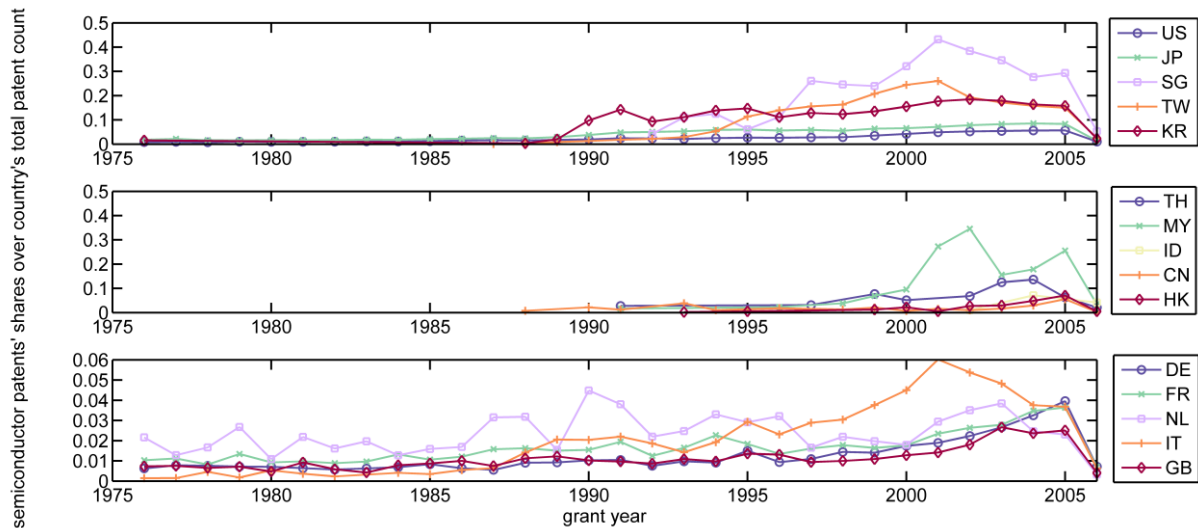
We have seen how a few latecomer countries (S.Korea, Taiwan and Singapore) have successfully narrowed their gap in terms of patent shares in the semiconductor while others, notably China, Hong Kong, Malaysia, Thailand and Indonesia, have failed to do so, at least until 2006. In the rest of the thesis, we will study in details which inventive strategies have been followed by Korean, Taiwanese and Singaporean firms. Here, we briefly investigate whether successful and unsuccessful catching-up countries started from similar conditions and if and how they differed in their sources and speed of technical improvements.

### *2.3.1 Similarities and differences of initial conditions*

Since, in this thesis, we analyse catching-up in the Semiconductor industry, we first need to ask ourselves whether successful catching-up countries were allocating more inventive effort to this industry compared to their competitors, before their catching-up process took off. Figure 5 shows how, until the mid-1990s, Taiwan's and Singapore's focus on Semiconductors was not higher than in the unsuccessful catching-up countries. It was also comparable to US, Japan and European levels (note that the latter are reported on a different scale). To the contrary, Korea's focus on the Semiconductor industry was preceding the take-off of its world's patent shares. This might signal a pre-existing comparative advantage due to a specific targeting of Semiconductor technologies. Once the catching-up process took-off, however, the importance of Semiconductors within each of the successful latecomers rapidly increased. At the beginning of the 2000s, Semiconductors patents accounted for around forty-percent of patent in Singapore and around twenty-percent in Korea and Taiwan. Similarly, in the 2000s, the share of semiconductor in a country's patent output increased for Malaysia and Thailand as well, fluctuating between twenty- and thirty-five percent for the former and around ten-percent for the latter. It remained much lower in China, Hong Kong and Indonesia.

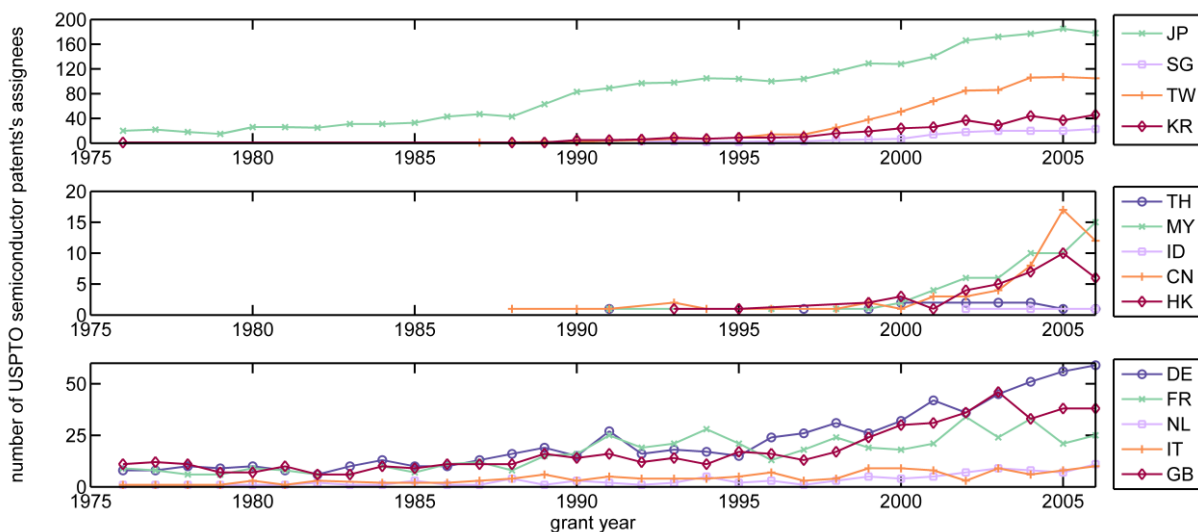
---

<sup>6</sup> Magee estimated that two-thirds of the total technological progress in information storage in the semiconductors is due to the use of better materials (Magee, 2010). Advances in industrial lasers and optical lithography also helped overcome the physical problems and the complexity of miniaturization.



**Figure 5: Semiconductor patent shares over country's total patent count**

Another key dimension of comparison is the number of organizations active in Semiconductor patenting in each country. This is reported in Figure 6<sup>7</sup>. In this case, we can see that all latecomers, regardless of their success, had only a few patent assignees until 1997. Then, more and more organizations started patenting in Korea, Taiwan and Singapore. About 100 different assignees had received at least one patent granted in 2006 in Taiwan, about 40 in Korea and 20 in Singapore. China, Hong Kong and Malaysia also experienced growth in the number of patenting organizations, even though of a smaller magnitude. China had 13 of them in 2006 (after reaching 17 assignees in 2005), Malaysia 15 and Hong Kong 5.



**Figure 6: Number of organizations active in Semiconductor patenting by country**

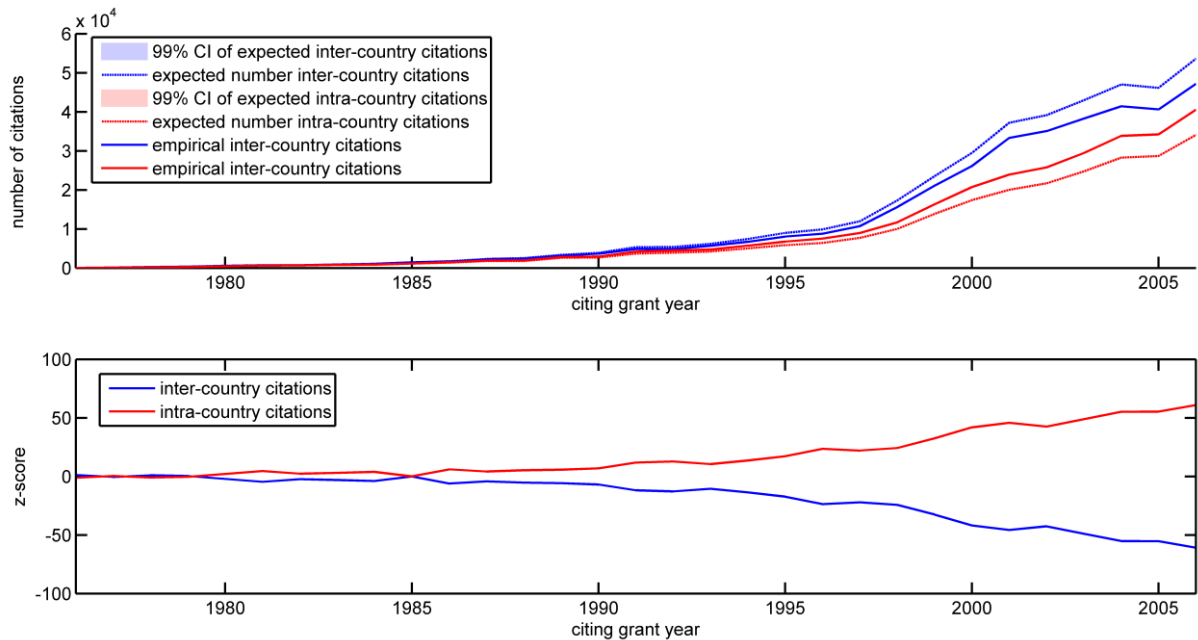
<sup>7</sup> We use harmonized assignee names included in the 2006 version of the NBER-USPTO database (Hall et al, 2001). We further cleaned typos in their names manually. Note that no effort was made to harmonize subsidiaries below the same holding. This is done on purpose, as different subsidiaries can follow very different inventive strategies and have different technical capabilities.

Obviously, the geographic and economic size of these countries differ widely, this may partly explain differences in the number of patenting organizations. The appropriate way to compare these countries is to look at the strength of their national semiconductor industry. If we take the world patent shares, the national focus on semiconductors and the first year of patenting as a reference, we can conclude that successful and unsuccessful catching-up countries shared relatively similar initial conditions in terms of number of inventive agents, strength of technical capabilities in the semiconductors and national targeting of semiconductor technologies, from the time of appearance of the first patenting activity in Semiconductor technologies until around 1995. On the other hand, there are two sources of dissimilarities. First, as we discussed, since the very beginning Korea showed stronger national focus on the semiconductors. Second, despite relatively similar numbers of patenting organizations in the mid-1990s, their type is markedly different. Patenting organizations from Korea and Taiwan were national firms. The largest assignees in Korea, at least until 2006, were Samsung, Hyundai Electronics, Hynix and LG Semiconductors. In Taiwan they were, Taiwan Semiconductor Manufacturing Corporation, United Microelectronic Corporation, the Industrial Technology Research Institute, Vanguard and Winbond. These are all local organizations. Singapore had a more mixed model. Its largest patent assignees up to 2006 were Chartered, a local foundry, and ASTAR, a government agency, but also foreign multinationals like Micron Technologies, ST Microelectronics and Texas Instrument. A similar strategy has been followed by Hong Kong, with a mix of local (ASAT and the Hong Kong University of Science and Technology) and foreign (Chartered, Motorola and Freescale) assignees. In contrast, Malaysia, Indonesia and China almost only had foreign companies in the list of their top assignees up to 2006.

### 2.3.2 *Sources of technical improvements*

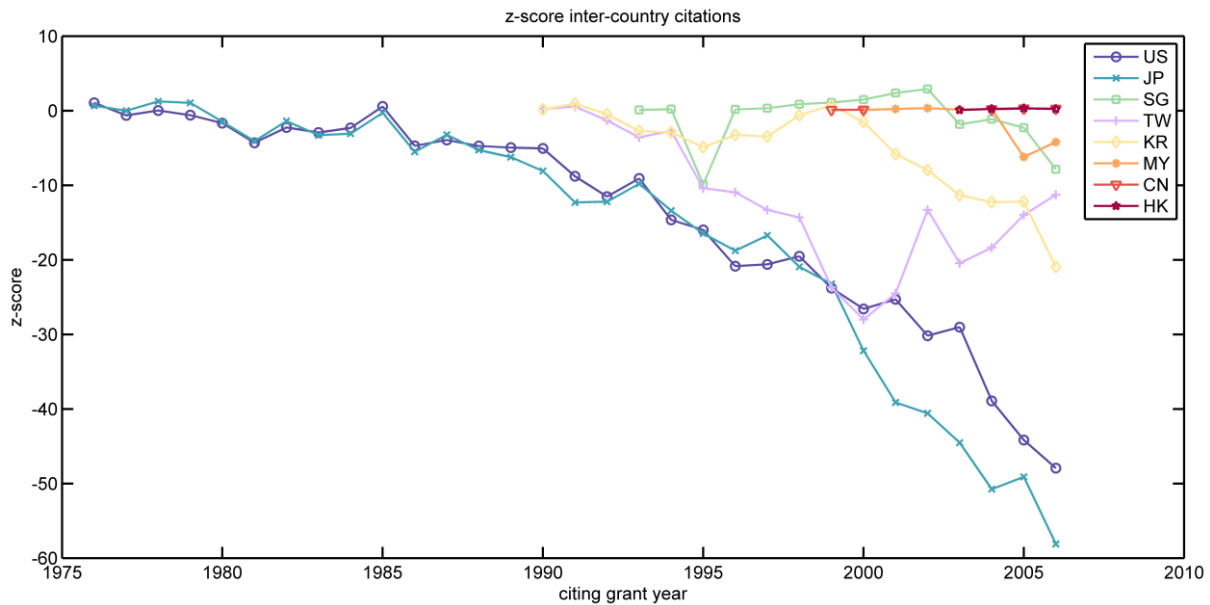
Given that, as we have argued, the improvement of technical capabilities is a key driver of long term market success, two additional key dimensions of comparison between successful and unsuccessful catching-up countries are the sources and speed of technical improvements. We measure them by looking at patent citations. Patents disclose the technical description of an improved technology. We interpret the existence of a citation between patents as a measure of the technical relatedness between the improved (i.e. the cited) and the improving (i.e. the citing) technology, rather than speculating on the existence of an intentional knowledge flow between the two inventions. This is because, as discussed by Criscuolo and Verspagen (2008) and Alcácer and Gittelman (2006) many citations are added by the patent examiners rather than the applicant, during the examination process. We first look at the geographic origin of the improved technologies, i.e. at the sources of technical improvements. The inclusive growth of the semiconductor industry and the globalization of its inventive activities would lead one to think that technological improvements have a more and more international character. In other words, one could think that it would be more likely to observe inventions improving on technologies that have been invented elsewhere. However, this is not the case. Indeed, the opposite is true. Semiconductor related inventions are more and more likely to improve upon locally invented technologies *than expected by chance*. To measure that, we computed the expected number of between-country and within-country citations, and their standard deviation, if one would randomly swap citations while preserving the number of patents granted to each organization and country, the number of citations made and received by each patent and their time structure (i.e. the difference between the citing and the cited grant year). We then computed z-scores by subtracting the expected number from the observed one and dividing by the standard deviation. The z-

score then indicates whether intra- or inter-country citations are higher, equal or lower than expected. Note that, to preserve the characteristic preference for self-citations, we have only randomized inter-assignee citations. More information on this randomization process can be found in the Appendix A.2.1. Figure 7 shows the results.



**Figure 7: The increasing reliance on national sources of technical improvements**

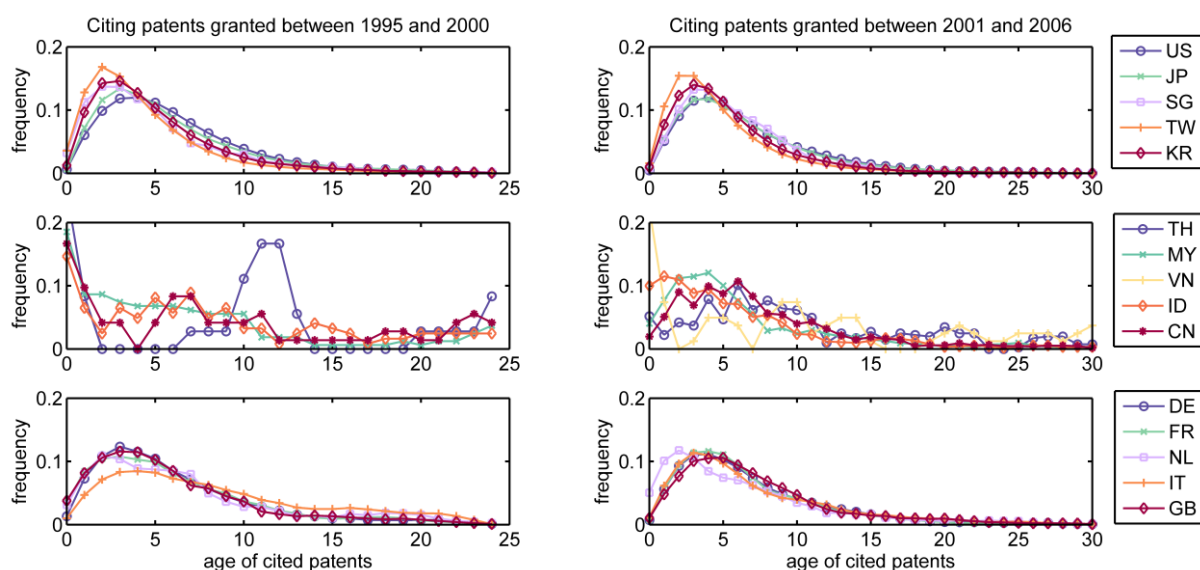
Approximately since the rise of Japan, and much more strongly after the entrance of the other Asian latecomers, there are more and more intra-country citations than expected by chance. This means that, when they were not improving upon their own technologies, new entrants, progressively increased their preference for improving upon inventions that were developed by another organization from the same country, rather than foreign ones. This might sound at odds with the globalization narrative. It is not. It is actually a sign of the effectiveness of the globalization of inventive activities. As we have seen in Figure 6, despite the rapid growth, there are still much less patenting organizations from latecomer countries than from US and Japan. Therefore, when citations made by patents granted to assignees from Korea, Taiwan, Singapore and the unsuccessful catching-up countries are randomly reshuffled, they have a small probability to be reconnected to assignees from the same countries. Therefore, the expectation is very low. The fact that there are indeed many more intra-country and between-firms citations than expected clearly shows that local technical capabilities have improved to the point that locally-invented technologies attract more attention from national organizations different from the assignee than foreign invented ones do. Is this true to the same extent across leaders, followers, and successful and unsuccessful catching-up countries? Figure 8 provides the answer by disaggregating the z-scores at the country level. For convenience, we only report the z-score for inter-country citations. US and Japan clearly became increasingly inward oriented, as there are less and less between-firms and between-countries citations than expected by chance. Similarly, Korea and Taiwan also progressively developed a preference for improving upon locally invented technologies. Interestingly, this is not the case for Singapore, China, Malaysia and Hong Kong. This shows how patenting organizations in these countries have still preferred to improve upon foreign technologies, and neglect local sources of technical improvements.



**Figure 8: Industry-building by improving upon local technologies**

### 2.3.3 Speed of technical improvements

Inventions typically improve upon existing technologies. By looking at the difference between the grant year of the citing and the cited patents, we can determine how old the sources of technical improvements are. Disaggregating these trends by countries reveal another key difference between successful and unsuccessful catching-up countries in the semiconductors. In Figure 9, we show the probability of citations to patents of different ages, computed for each country. In the left column, we included only citing patents granted between 1995 and 2000. In the second one we include citing patents granted between 2001 and 2006. The shape of these curves is fairly typical across technologies, as shown by Jaffe and Trajtenberg (2002) and Hall and colleagues (2001). New technologies need time to be recognized and the improvement of old technologies tend to attract much less inventive effort. This causes the particular shape of the probability density function. The comparison by country is insightful. In both periods, Korean, Taiwanese and Singaporean patenting organizations had a stronger preference for improving upon recently improved technologies, compared to firms from the other countries. In contrast, patenting organizations from unsuccessful catching-up countries do not seem to have any clear preference. This could be because probability density functions are noisy when there are few observations. This is the case for unsuccessful catching-up countries, which, having fewer patents, also have fewer backward citations. To reduce this problem we computed the cumulative distribution function (cdf). We report the results in the Appendix A.2.2. Using the cdf, we have also performed a Kolmogorov-Smirnov and an Anderson-Darling test of equality in the distributions between countries. Both tests confirm that the distributions are statistically different. This showed a statistical preference for patenting organizations from successful latecomers in the Semiconductors to perform quick cycles of improvements of existing technologies.



**Figure 9: Fast improvement of semiconductor technologies by successful catching-up countries**

## 2.4 Conclusions

Semiconductors are a fascinating industry for an economist interested in studying the interplay between technological change, inventive strategies and catching-up. Given its central role as engine of economic growth, and the occurrence of both successful and unsuccessful catching-up ventures, it is very insightful to analyse the technological and industrial dynamics in this industry and investigate what made successful catching up organizations special. In this chapter, we have provided a preliminary overview on that.

We have shown that inventive activities in the Semiconductor industry became more and more globalized than expected by looking at the overall technologies patented at the USPTO starting from the mid-1990s. This corresponds with the prominent entrance of patenting organizations from Korea, Taiwan and Singapore. Firms from these countries managed to rapidly gain patent shares, mostly by eroding technological competitiveness of Japanese firms. The increase in invention output was closely matched by a fast increase in market shares. At the same time, the Asia-Pacific region became the most important market in the world for Semiconductors, as measured by dollar amounts of billings. We have also discussed how the technology dynamics within the industry clearly favoured the entrance of new players. In particular, the increase in technical modularity fragmented the value-chain of semiconductor devices and opened entry opportunities for specialized players. However, not all latecomers were able to effectively take advantage of these opportunities. A few countries, notably China, Hong Kong, Malaysia, Indonesia and Thailand, which started patenting at the USPTO and manufacturing semiconductors approximately at the same time of Korea, Taiwan and Singapore, failed to narrow the gap in terms of patent shares. These countries started from relatively similar initial conditions (at the time of their first USPTO patents) in terms of national inventive focus on semiconductors and number of patenting organizations. However, a number of clear dissimilarities emerged between successful and unsuccessful catching-up countries. First, the former had a rapid expansion of the number of patenting organizations within their borders. This shows that the number of active agents within the innovation system increased. Second, the composition of these agents was very different across successful and unsuccessful catching-up countries. Patenting organizations in Korea and Taiwan were mostly local companies. Singapore and Hong Kong had a mix of local firms and for-



eign multinational corporations. China, Malaysia, Indonesia and Thailand inventive agents were mostly foreign companies, at least until 2006. Third, in Korea and Taiwan, more than in any other catching-up country, patenting organizations progressively developed a strong preference for improving locally invented technology. Fourth, patenting organizations in successful catching-up countries had a stronger focus on rapid cycles of technological improvements than leader and unsuccessful catching-up countries. A similar conclusion has been recently discussed by Lee (2014), who showed how Korean and Taiwanese firms preferentially entered industries with short life cycles. Taken together, these findings suggest that successful catching-up countries were actively engaged in industry-building activities, more than what could be observed in other countries that shared relatively similar initial conditions at the time of their first USPTO semiconductor patents. They had a growing number of national patenting organizations; they rapidly shifted from improving foreign technologies to focus on national technologies and they went through rapid cycles of technological improvements.

These findings shed some light on the technological path followed by firms from successful catching-up countries. They also highlight how their inventive strategies might have been markedly different from those adopted by patenting organizations from the US, Japan and unsuccessful catching-up countries. In the next chapters, we will analyse in detail the inventive trajectories that have been followed by these firms across the technology space. We will also reveal if they successfully upgraded their comparative advantage along technology life cycle stages.



# 3 ENGINEERING TRAJECTORIES, RANKING OF DESIGN PROBLEMS AND CATCHING-UP

## ABSTRACT

Technology is a complex adaptive system whose direction of development depends on the interdependences of engineering problems, on the propagation of technical solutions and on firms' strategic decisions. We argue that the complex interactions determining the direction of change at the technology level can be expressed in terms of two key dimensions: the presence or lack of persistence in the *ranking of the system of engineering problems* and the presence or lack of *stability in the approaches to problem solving*. These dimensions are the projection in the engineering design space of technological trajectories defined at the product level. We design a set of metrics to identify changes in these two dimensions, and disclose statistically significant path-followers, path-changers, ranking-followers and ranking-changers. We analyse the global Semiconductor industry. We identify two discontinuities in the direction of technological change. In the first half of the 1990s, a disruption in the main approaches to problem solving occurred because of the introduction of new semiconductor device structures. In the first half of the 2000s, we observed a change in engineering trajectories, coupled with a change in the ranking of problems, caused by advances in the miniaturization of semiconductor devices allowed by the diffusion of MOSFET technologies. We show that firms from latecomer countries were in general significant path- and ranking-followers in the 1990s. This partially changed in the first half of the 2000s, when some of them became significant ranking-changers.

### 3.1 Introduction

Technology is in continuous evolution. Depending on the direction of technological change, capabilities can be reinforced or outdated. This affects the catching-up fortunes of new entrants and the resilience of incumbents. Studies on catching-up emphasized the characteristics of sectoral technological regimes as key determinants of the ease of catching-up (Lee, 2014; Lee and Lim, 2001; Breschi et al. 2000, Malerba and Orsenigo, 1997, Klepper, 1997). These are defined as the peculiar features of the set of technologies underlying a given industry. Malerba and Orsenigo (1997) empirically showed that the level of concentration of economic activities, entry of new players and the stability of the ranking of core innovators, are strongly affected by the availability of technological opportunities, the cumulativeness of technological improvements and the ease of appropriating returns to innovation. Lee and Lim (2001) highlighted that, with respect to catching-up, the stability or discontinuity in the technological trajectory plays a particularly crucial role. The authors define three types of catching-up strategies. Firms that focus their learning efforts on reducing the technological gap by replicating what the current leaders have been doing are defined as path-followers. In contrast, the term path-creator refers to firms that, while learning existing technologies, improve them to a point that they become fundamentally different from what they used to be. Finally, Lee and Lim also defined path-skippers. These firms jump over generations of technologies during their learning process and are able to engage quickly in frontier research, by skipping intermediate learning stages. This classification is not exclusive for catching-up firms. If we assume that a set of possible technical paths exist in an industry at any point in time, and we broaden our focus on all firms populating an industry, we can use a similar classification for any firm. Firms, be they large or small, incumbents or new entrants, can be path-changers or path-followers, in a similar vein as they can be exploiters of cumulative knowledge or explorers of new research directions. These strategies can be reactive or proactive to change, might be explicitly defined within the firm, or just be the result of random events that are implicit in the research and development process. They might also have alternative fortunes depending on the configuration of the system of technologies at a given point in time. Yet, firms' location with respect to the path of technical improvements only refers to one dimension of technological change, namely the presence or lack of stability in the *approaches to engineering problem solving*. We argue that the direction of technological change has a second crucial dimension: the presence or lack of stability in the *ranking of engineering problems*.

In this paper, we use patent citation data to trace the direction of technological change along these two fundamental dimensions, the ranking of engineering problems and the stability of problem solving approaches. By doing so, we reconstruct the engineering trajectories followed by an industry over time. We apply the *network of main paths* (NMPs) methodological approach for these purposes. The NMPs is a methodology that helps disentangling the complexity of large patent citation networks. It builds on the work by Hammond and Doreian (1989), Garfield et al. (1964) and Garfield (1979), and has been extended by a branch of research which have applied it to the definition and analysis of technological trajectories in several industries (Verspagen, 2007, Fontana et al., 2009, Martinelli, 2008; 2009 and Bekkers and Martinelli, 2010). This line of research provides an extremely promising way to look at patent data and technological performance of firms. The NMPs reveals much richer insights than simple patent or citations counts. We contribute to this methodology by defining a set of indices that reveals the presence or lack of stability in the ranking of the systems of engineering problems and approaches to problem solving. We also define a method to assess whether players are statistically significant path-followers or changers and ranking-followers or changers. Finally, we design an

unbiased measure of firms' centrality on the main paths of engineering improvements, which corrects for network characteristics, time of entry and firm size. We use the NBER patent citations data file (Hall et al., 2001), which contains detailed metadata for patents issued by the USPTO between 1976 and 2006.

We analyse the Semiconductor industry as a case study. Semiconductors provide the perfect ground to study the interplay between the direction of technological change, players' strategies and catching-up. The industry has been once dominated by US firms. Then, as shown in Chapter 2, about four decades ago, Japan has risen as a major player, followed in the 1990s by South Korea, Taiwan and, partly, Singapore. The success of these latecomers has been striking. In this work, we investigate which strategies have been adopted by latecomers and incumbents to guide their search for high-quality engineering solutions in the technology space. We also assess the effectiveness of these strategies in building technological capabilities, measured as a firm's ability to affect the direction of technological change. The paper is structured as follows. In Section 3.2, we introduce the theoretical framework that guides our analysis and define the two key dimensions of the direction of technological change: the ranking of engineering problems and the stability of approaches to problem solving. Section 3.3 explains the methodological approach and introduces the set of indices used to analyse these two dimensions and players' strategies. Section 3.4 presents the data used. Section 3.5 reports our findings for the analysis of the direction of technological change in the Semiconductors. Section 3.6 reveals players' contribution to technological change. Finally, Section 3.7 discusses the effectiveness of players' strategies in terms of actively influencing the direction of technological change.

## 3.2 Theoretical framework

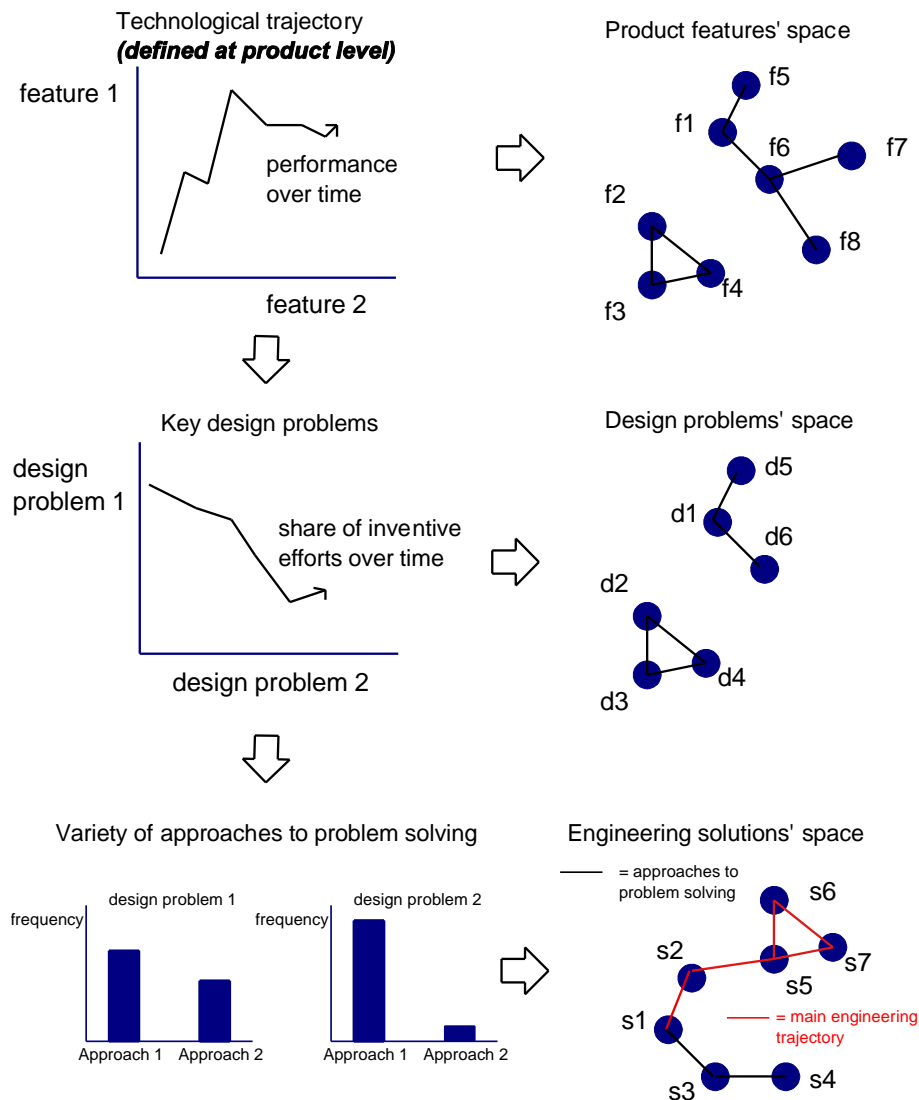
The evolution of technology determines and constrains entry, catching-up and growth opportunities (Klepper, 1997, Breschi et al. 2000, Lee and Lim, 2001). Malerba and Orsenigo (1997), theoretically defined, and empirically showed, how technological regimes determine ease of entrance, stability of the ranking of innovators and concentration of inventive activities. The authors define technological regimes as particular combinations of some fundamental properties of technologies. These are opportunity and appropriability conditions, degree of cumulativeness of technological knowledge and characteristics of the knowledge base. Lee and Lim (2001), conjecture that, as far as catching-up is concerned, cumulativeness of technical advance and the predictability of the technological trajectory are the most important dimensions of the technological regime of an industry. They argue that large knowledge cumulativeness requires experience to succeed in the current technological search, reducing success chances for new entrants. On the other hand, they hypothesize that stable technological developments helps latecomers fixing their R&D target. Lee and Lim, define technological trajectories as changes in the vector of technical features along the sequence of product generations. Knowledge cumulativeness refers to how much working experience in the previous generation help in mastering the subsequent generation. They highlight how stability in the trajectory and knowledge cumulativeness does not necessarily co-appear. They use the dynamic random access memory (DRAM) industry as an example of a case where the future characteristics of the technology are known well in advance but require developing a new set of knowledge. They argue that this scenario helps explaining the successful entrance of Korean memory makers. Lee and Lim's definition of the technological trajectory builds on the seminal work by Dosi (1982). Dosi theoretically defined the direction of technological change as emerging from the interaction between technological paradigms and technological trajectories. Dosi defined the former as "... [a] 'model' and a 'pattern' of solution of

*selected technological problems, based on selected principles derived from natural sciences and on selected material technologies*" (Dosi 1982). A technological trajectory is then defined by Dosi as "...the 'normal' problem solving activity determined by a paradigm. [It] can be represented by the movement of multi-dimensional trade-offs among the technological variables which the paradigm defines as relevant..." (Dosi 1982). Within the same paradigm, firms can explore different strategies. Therefore, several technological trajectories can co-exist. The same is argued by Malerba et al. (1999), who describe trajectories as the outcome of firm's choices between competing alternatives, like cost and performance of a given product.

This shows how technological trajectories are usually understood and defined, in the innovation literature, at the product-level. Yet this is not necessarily the best level to study the co-evolution of technology, industrial dynamics and catching-up. Products are collection of components and technologies (Murmann and Frenken, 2006; Arthur, 2009). It is precisely the structure of the system of technologies that determines constraints and opportunities for improvements along the spectrum of features that stand behind a product. Furthermore, learning is a problem-solving activity. As such, it happens at the technology level, as it is at this level that engineering problems arise. Therefore, we argue that, in order to understand the scale of change in capabilities and, consequently, the availability of entry and catching-up opportunities, technological trajectories need to be studied at the level of technology domains rather than at the product level. More precisely, we argue that the direction of change at the technology level can be expressed in terms of two dimensions, namely the *ranking of engineering problems* and the *stability of problem-solving approaches*. Figure 10 illustrates this argument.

As discussed in the literature (Malerba et al., 1999; Dosi, 1982), technological trajectories are intended as the sequence of design choices made by firms when they face trade-offs between the features of a given product that they want to improve. Obviously, some features may be complementary and can be simultaneously improved; some others may be in contrast. Potentially, one could trace performance improvements over time between complementary and competing product features, as sketched in the upper-left corner of Figure 10, and done by Malerba et al. (1999). For instance, thanks to miniaturization, integrated circuits can be made smaller and more powerful at the same time. In contrast, increasing their computing power while simultaneously reducing energy consumption or heat dispersion is a much more challenging task. This generates implicit relations between product features, which could in principle be described by a product features' space, where features that can be simultaneously improved would be linked together, as illustrated in the top right corner of Figure 10. The breadth of features of a given product, its functional performance and the ease of breaking trade-offs between features, crucially depend on generating innovative solutions to engineering challenges. Therefore, the choice of which product features one would like to improve simultaneously determines the distribution of inventive efforts across a given set of engineering design problems. Firms might have different beliefs on which product features should be improved and which engineering challenges have to be tackled. It follows that the aggregation of firms' inventive efforts determines a *ranking of engineering problems*, based on the amount of collective efforts that they attract. Therefore, one could also assess the evolution of such ranking by tracing the share of inventive efforts devoted to the search of solutions to each problem, as fictitiously illustrated in the mid-left panel of Figure 10. Similarly to what described for product features, design problems are also related to each other's. Solving a given design challenge may allow finding a solution to related ones. In other words, the space of engineering problems has a latent structure, which is deeply related to the structure of the product feature space. Again, such structure could

be represented as a network of engineering problems linked by a relationship similarity in possible solutions.



**Figure 10: The relationship between technological trajectories, ranking of design problems and engineering design trajectories**

Because of the uncertainty intrinsic to the innovation process, solutions to design problems may be searched in multiple ways. Different approaches may be followed with different frequencies. Despite the variety of possible approaches, problem solving tends to be path dependent. This is because agents (be they inventors or firms) tend search for solutions to problems by leveraging on knowledge that they have learned in the past through individual and collective effort. This generates paths of incremental technical improvements that shares similarity in problem-solving approaches. Such paths are sketched in the bottom-right panel of Figure 10, where nodes represent solutions linked by similarity in problem-solving approaches. Sometimes, new approaches are pursued by explorative agents. If successful, they effectively set new paths that others can follow. The stability of problem-solving approaches refers to the frequency and the extent that new paths are created. Therefore, the ranking of engineering problems and the stability in problem solving approaches are the projection in the engineering design space of technological trajectories defined by decisions on how to tackle trade-offs at the product level. For this rea-

son, we refer to the paths of technical improvements as *engineering trajectories*. Paths that keeps attracting more and more inventive effort, implicitly reveal which are the prevailing problem-solving approaches followed by inventors in a given industry. Therefore, they can be defined as the main engineering design trajectories historically followed in an industry. In our toy network sketched in the bottom-right panel of Figure 10, the main path is highlighted in red.

The relationship between technological and engineering trajectories is bidirectional. Product design choices, which define technological trajectories, depend on the co-evolution of technical progress, market needs and firms strategies. In this work, we focus on the first factor of influence, namely, technical progress. Solutions to design challenges define the direction of technical progress at the product level and expand the variety of possible applications of a given technology. They also enlarge the set of available features of related products. Therefore, engineering trajectories, defined at the technology level, determine which technological trajectories at the product level are possible. This is known in the literature as the technology-push hypothesis. Of course, the opposite direction of causality can also exist. Market demand for certain applications and features creates incentives for investing resources to solve particular engineering problems that characterize them. This is referred to as the demand-pull hypothesis. What is sometimes neglected in the technology-push and demand-pull debate is the way change propagates. Problems are interconnected. Solutions to technical bottlenecks spark the emergence of new applications, which ultimately define new technical challenges. For instance, miniaturizing microchips allows a variety of applications in electronics to emerge. However further miniaturization defines a set of key technical challenges. For example, as the technology scaling reaches channel lengths less than a micron, second order effects, which were ignored in devices with longer channel length, become very important. Some of these effects are velocity saturation and degradation due to overheat. Solutions to these problems propagate to other technologies. This discloses the existence of a system of technology. Finding a solution to these semiconductor-related technical challenges made it possible to have lighter and more powerful computers, or flatter and brighter television and computer screens. It also allowed creating new products, like smartphones and tablets. These new products ultimately define new technical challenges. Heat dispersion and energy consumption are problems that are much more important in laptops, tablets and smartphones than in desktop personal computers. Similarly, better brightness, image sharpness and energy saving features became available for modern screens of TV and portable devices because of the miniaturization of semiconductors. However, they also posed different engineering problems compared to old generations of these technologies. Therefore, the connections of the system of engineering problems can change over time. There might always be a central problem (like the miniaturization of microchips) but the related problems might change, depending on changes in the technological trajectory at the product level.

Ultimately, the ranking of engineering problems and the stability of problem-solving approaches determine the scale of change in the required technical capabilities. These changes affect the success of new ventures and the survival of old ones. This is why analysing the evolution of these two dimensions is critical for our understanding of catching-up. As Lee and Lim (2001) pointed out, changes in trajectories create entry opportunities but also increase risk because of the intrinsic uncertainty related to exploring new paths. Similarly, changes in the ranking of engineering problems can potentially create discontinuities in the otherwise cumulative nature of knowledge generation, making prior experience less important. These changes are the results of firms' strategies. Therefore, based on the two dimensions of technical change discussed here, we can categorize firms as path-changers or followers and as ranking-changers or followers. We define path-followers as agents that focus on incremental innovation by applying well-



established approaches to problem solving. In contrast, path-changers attempt to explore new paths. This can potentially lead them to become vehicles of radical change. Ranking-followers, focuses their innovative effort on seeking solution to engineering problems that are currently perceived by the majority as the most important ones. Ranking-changers are relatively more focused on problems that comparatively attract less attention. This could be because they are related to niche products or because they are overlooked by the majority. Yet niches can develop into large markets depending on the evolution of technology and demand. Similarly, problems that are overlooked today can become very central in the future.

Now, the question is, *how can we empirically trace the evolution of engineering problems and approaches to problem solving?* Trajectories of technical improvements can be detected by tracing paths of patent citations, as done in Verspagen (2007), Fontana et al. (2009), Martinelli (2008 and 2009) and Bekkers and Martinelli (2010). This strategy also makes the identification of trajectories empirically possible. To trace technological trajectories defined at the product level one would need a great wealth of multidimensional data on functional performance and cost, to be able to identify how firms navigated along trade-offs. Studying engineering trajectories at the technology level can be done by using publicly available patent data. Patents are proofs that novel and non-obvious solutions to given engineering problems have been found. A few characteristics of the patent systems can be exploited to identify paths of engineering improvements, and disclose the system of engineering problems and the approaches to problem solving used<sup>8</sup>. A patent is examined by experts in the subject matter *at its claim level*. Claims are classified into subjects and searched for existing prior art. The goal of this search is to find evidence that what has been claimed is entirely or partly lacks novelty. Such evidence is then used to narrow the temporary monopolistic legal protection granted by the patent to what is *truly new*. As stated in the USPTO Manual of Patent Examination Procedure (MPEP)<sup>9</sup>, any document used in the rejection of a claim is used as a reference. This means that patent citations effectively highlight the existing prior art of an invention. The definition of what constitutes prior art in the USPTO system, clarifies why we claim that citations highlight similarities in the approach to problem-solving. In this respect, the MPEP states the following. *“During patent examination the claims are given the broadest reasonable interpretation consistent with the specification”* (USPTO, 2014, MPEP Section 904.01). *“Not only must the art be searched within which the invention claimed is classifiable, but also all analogous arts must be searched regardless of where the claimed invention is classifiable. The determination of what arts are analogous to a particular claimed invention [...] depends upon the necessary essential function or utility of the subject matter covered by the claims, and not upon what it is called by the applicant. For example, for search purposes, a tea mixer and a concrete mixer may both be regarded as relating to the mixing art, this being the necessary function of each”* (USPTO, 2014, MPEP Section 904.01c).

Therefore, paths of citations between patents disclose sequences of improvements of existing prior art. As such, they reveal commonalities in the approaches to problem solving. These paths do not have to be necessarily intended as conscious use of prior knowledge. In contrast, the addition of citations by examiners allows exploiting expert opinions on what constitutes re-

---

<sup>8</sup> In this work we use US patents. Therefore, the characteristics of the patent systems that we discuss here refer to the functioning of the system run by the United States Patent and Trademark Office (USPTO). Differences in the way inventions are defined and classified exists with other major patent offices in the world, such as the European Patent Office (EPO) and the Japanese one.

<sup>9</sup> The USPTO MPEP is available online at this website: <http://www.uspto.gov/web/offices/pac/mpep/>

latedness of approaches. Therefore, paths of citations can be legitimately interpreted in terms of similarity of approaches to problem solving. This means that changes in the composition of the relevant prior art over time reveal changes in the trajectories followed to tackling engineering problems. Furthermore, changes in the propensity of some paths of citations to attract further inventive efforts reveal changes in the ranking of the system of engineering problems. This theoretical argument stands behind the methodological approach used in this work. We introduce this approach in the next section.

### 3.3 Methodology

To make our theoretical framework operational we need to perform a number of tasks that will ultimately allow us to investigate if the process of technological change in the Semiconductor industry followed a cumulative or disruptive path. We will then be able to identify path- and ranking-changers in the industry. The tasks needed to accomplish those goals are listed in Table 1, together with the methodology used to tackle them. In the following subsections, we explain the methodology that stands behind each task in details.

**Table 1: Methodological framework roadmap**

| Task  | Methodology/Measure used  |
|---|---|
| Identification of technically influent patents and of the main paths of technical solutions | Network of Main Paths (NMPs)  |
| Measuring the ranking of the system of engineering problems                                 | Attractiveness of the largest and second largest components of the NMPs |
| Measuring the stability of problem-solving approaches                                       | Patent persistence within and between NMPs two largest components       |
| Measuring centrality on the paths of engineering improvements                               | Path-centrality index   |
| Measuring agents' contribution to the direction of technological change                     | Path-changing and ranking-changing indices                              |

#### 3.3.1 Identification of technically influent patents and of the main paths of technical solutions

Patent technical and economic value is highly skewed (Silverberg and Verspagen, 2007). Therefore, the merit of sampling important patents out of the whole population is twofold. First, we reduce noise in the citation network generated by patents of dubious technical and economic value. Second, we reduce a large-scale citation network to its backbone. This allows the fundamental structure of the network to emerge. We define technical importance as a patent's ability to influence technical change. Therefore, technically important patents are those laying on the most central paths of engineering improvements. To identify those paths we rely on the *Network of Main Paths* approach (NMPs). The first step of the methodology consists in assigning weights to each citation. To do that we rely on the Search Path Node Pair (SPNP) algorithm as explained in Batagelj (2003) and De Nooy et al. (2005). The SPNP algorithm counts the number of times that each node lays on all possible paths that connects any node to anyone else. As such, it has a very clear interpretation as a measure of connectivity. When normalized over the total number

of possible paths, the SPNP value of node  $i$ , tells us the probability that a random walk (i.e. following any path regardless of its length) between any pair of nodes in the network passes through node  $i$ . The flow-normalized SPNP is analogous to the measure of random walk-betweenness centrality (RWBC) proposed by Newman (2005). However, the former anticipated the latter, as was first discussed in Hummon and Doreian (1989). Both SPNP and RWBC are calculated taking into account all possible paths between any pair of nodes. This distinguishes them from the standard measure of betweenness, which takes into account only the shortest path(s). We argue that this is a better measure of centrality in the context of technological trajectories as if one would assess the technological relatedness of a reachable pair of patents by looking only at the shortest path of citations it would exclude alternative sequences of technological improvements which connect the same pair of patents. These alternative sequences can be very informative as, for instance, they might pass through a different technological area or involve different companies. Once citation weights have been calculated, the original network is reduced to its main paths by an iterative process. The algorithm starts from source nodes (patents that are cited but do not cite anyone) and identifies the main paths of citations by following at each junction the citation that carries the highest weight, until a sink node (a patent that cites someone but is not cited) is reached. The procedure is explained with the help of a fictitious network shown in Figure 11.

**Figure 11: A dummy citation network**

The dummy network is made of 22 patents. The SPNP weight for every citation is shown above each line. Citations are represented by arrows, whose head indicates the direction of improvement (technology disclosed in cited patents is improved by citing ones). Consider, for instance, the citation between patents 5 and 9. It has a weight of 16. This is given by the multiplication of the number of patents reaching patent 5, plus 5 itself (i.e. patents 1 and 5), and the number of patents reached by patent 9, plus 9 itself (i.e. patents 9, 13, 15, 17, 14, 16, 18 and 19). To identify the NMPs we start from the set of source nodes (patents 1, 2, 3 and 4) and follow at each step the citation carrying the highest weight, till one of the end nodes is reached (patents 17, 18, 19, 21 and 22). By repeating this procedure for each start point we identify the NMPs, which, in the example above, is made of two components whose nodes are coloured in black (main one) and grey (second one). It is important to notice that the two components of the NMPs are not separated if we look at the original network, but the white nodes that connect them have a negligible importance from the point of view of the paths of engineering improvements. This distinction let different *systems of engineering problems* to emerge. Patents within components are mutually related by paths of important improvements that address engineering problems connected to

each other's. If the component is very large, the relationship can be loose. Nevertheless, it exists. In contrast, solutions disclosed in patents that are located in different components are not related. Therefore, a dynamic analysis of the NMPs allows detecting changes in the composition and ranking of the systems of engineering problems and tracing the stability of problem-solving approaches. The dynamic approach consists of cumulating networks at different points in time (e.g. from time  $t$  until  $t+1$ , then from  $t$  until  $t+2$ , and so on). This allows assessing how the entrance of newly granted patents in the system affects the presence of old ones in the network of main paths and the size ranking of its components. When newly granted patents largely connect to the main component of the NMPs the ranking of the systems of problems is preserved. When the relative majority connects to any other component, there is a change in ranking. A new system of engineering problems now attracts the bulk of innovative efforts. Newly granted patents also reveal which approach to problem solving has been followed. To understand how this is the case let us use once again the dummy network illustrated in Figure 11. Suppose that a set of 10 new patents would enter in the network at time  $t+1$ . For simplicity let us imagine that the 10 patents will connect (directly or indirectly) just to one endpoint<sup>10</sup>. Then, three cases might be observed.

- **CASE 1:** If the new entrant patents connect to patent 18 or 19 then the main component is still attracting the majority of innovative effort and all patents that were previously on the main paths within the largest component will still be found there. In this case, both the approaches to problem solving and the stability of the ranking of problems are preserved.
- **CASE 2:** New entrant patents connect to either patent 21 or patent 22. If that happens, the sequence of patents 3-7-44-20 becomes the root of the new largest component. In this case, the largest component is still the main attractor of innovative effort but now solutions to related engineering problems are seek through different approaches, as revealed by the fact that new patents connect to a previously less exploited prior-art.
- **CASE 3:** Let us now imagine that the new entrant patents will connect at time  $t+1$  to patent 17. In this case, what was formerly the second largest component becomes the main one and the ranking of the system of engineering problems is subverted.

### 3.3.2 *Measuring the ranking of the system of engineering problems and the stability of problem-solving approaches*

Obviously, in real-world networks the attachment of newly granted patents follows patterns that are more complex. Some of them might connect to the largest component and some to the second or any of the others. Yet, these patterns can always be reduced to two dimensions. From the point of view of the stability of the ranking of the system of engineering problems, what matters is whether the relative majority connects to a component different from the previously largest one. From the point of view of stability of approaches to problems solving, the extent to which new patents connect to previously unexploited prior-art, which could be either related to the main system of problems or not, indicates how much the search of solutions follows alternative approaches. Therefore, we only need to keep track of two measures, the *attractiveness* of the largest component of the NMPs, and the *persistence* of patents within and between the largest

---

<sup>10</sup> In reality, they could connect to any other patent as well, but this empirically happens with a probability inversely proportional to the time lag between the citing and the cited patent, for time lags longer than about 4.5 years, as shown by Jaffe and Trajtenberg (2002).

components of the NMPs over time. The *ranking* of a given system of engineering problems  $j$  (i.e. of a component of the NMPs) at time  $t$ , is measured by the attractiveness of the component of the NMPs that represents it.

$$attractiveness_{tj} = \frac{P_{tj}}{\sum_k P_{tk}} \quad (3.1)$$

Where the numerator  $P_{tj}$  is the number of important patents granted at time  $t$  that connect to component  $j$ . The denominator is simply the sum of all the important patents granted at time  $t$ , which connect to any of the component  $k$  of the NMPs.

The stability of approaches to problem-solving is measured by the following index, which quantifies persistence of patents within the components of the NMPs over time. The *patent persistence index* (PPI) for component  $j$  from period  $t$  to period  $t+1$  is equal to:

$$PPI_{j_{t \rightarrow t+1}} = \frac{N_j^t \cap N_j^{t+1}}{N_j^t} \quad (3.2)$$

Where the denominator  $N_j^t$  is equal to the number of patents in component  $j$  at time  $t$  and the numerator is equal to the intersection between the set of patents in component  $j$  at time  $t$  and the set in the same component at time  $t+1$ . Obviously, it is possible to use the same approach to quantify movements of patents across components. In this case, the numerator measures the intertemporal intersection between components. By cumulating NMPs from period to period, we are able to assess the stability of search strategies across the space of possible solutions to engineering problems. If newly granted patents in period  $t+1$  build on the same prior-art that was improved along paths defined in period  $t$ , then citations from the former will connect, directly or indirectly, to the latter. In this case, paths that were central in  $t$  remains central in  $t+1$ . This indicates stability in the approach to problem solving and in the corresponding technological trajectory. In contrast, when newly granted patents represent solutions found through searching alternative approaches, citations will point to a previously unexplored prior-art. This will disrupt the main paths, revealing a change in the approach to problem solving.

### 3.3.3 Measuring centrality on the paths of engineering improvements

We ultimately want to analyse catching-up strategies at the firm level. This requires defining a measure of the capability gap and identifies the strategies applied to reduce or close it. We define technical capabilities as the ability of affecting the direction of technological change. This ability is revealed by patent centrality on the NMPs. Companies whose patents are located in central junctions of the NMPs are able to influence the perception of which engineering problems are important and which approach to problem-solving is promising. Therefore, the effectiveness of technical catching-up strategies can be measured in terms of companies' ability to move to the centre of the system of engineering problems, i.e. to have patents centrally located in the NMPs. As we explained in Section 3.3.1, the SPNP index provides this information. However, the index suffers from a problem that hinders its comparability over time, which we need to solve before using it as a measure of company centrality on the paths of technical improvements. Two effects play a role here. First, let us remind the reader that the SPNP index is calculated at the node level. It counts the number of times that a given node shows up on the possible paths that connect any node to anyone else. Hence, if we sort the network topologically such that any

node is preceded by all nodes pointing to it, we immediately realize that nodes that stand in the middle will tend to get higher SPNP values. This is simply because, by construction, they have better chances to be found on several paths. The second effect comes from the empirical evidence on the growing size of patent citations networks. One has to consider that a growing number of newly granted patents per year implies a growing number of citations added to the network every year (unless the average number of backward citations per patent dramatically falls). Consequently, the number of possible paths between nodes increases (exponentially) with the size of the network. This implies that the probability of a node to lay on many connecting paths increases when the network grows and a giant component exists. These conditions are all fulfilled by our empirical network. Therefore, if we compare the centrality of a given company in the network of main path over time by using the average of the SPNP count of its patents, it is very likely to observe that the company becomes more and more central. This could be a “real” phenomenon or it might be purely because the SPNP count increases because of network growth. The SPNP therefore needs to be normalized. Dividing the SPNP count of every node by the total number of paths computed at every period of observation does not solve the problem because the probability of laying on many paths grows less quickly than the number of paths itself, hence creating another bias. To correct for these biases and allow us to compare SPNP weights over time, we need to rescale the SPNP weight measure. We call the new measure *Path Centrality Index* (PathC). The index for node  $i$  is calculated as follows.

$$PathC_i = \frac{SPNP_i}{\frac{1}{N_t} \sum_{i=1}^{N_t} SPNP_i} \quad (3.3)$$

PathC is the SPNP weight of node  $i$  divided by the average SPNP weight assigned to the subset  $N_t$  of patents granted in the same year  $t$  of patent  $i$ . Because of the very large number of patents per grant year (minimum  $N_t$  is around 5000), the average SPNP for any given grant year is a good approximation of the true expected SPNP value for a randomly chosen patent granted in the given year, assuming it would cite other patents randomly. Accordingly,  $PathC_i$  quantifies how good patent  $i$  has performed in terms of connectivity. It does so by comparing its performance with the connectivity achieved by other patents which entered in the network at the same time, therefore facing the same probability of getting high (or low) SPNP weights. Note that this does not mean that patent  $i$ 's performance will only be affected by patents granted in the same year, as the numerator of the fraction still depends on how important is patent  $i$  in terms of connectivity with respect to the whole network. This method is analogous to the fixed-effect approach proposed by Hall et al. (2001) to clean the number of forward citations received by a patent by any possible effect of network structure. As explained by the authors this simple method allows purging the data from systematic effects caused by change in the propensity to cite and in the number of citing patents. However, it also does wipe off any “real” technology-based effect. In particular, it does not allow us to compare absolute levels of centrality for the same agent (be it a patent or a firm) over time. Yet, it effectively serves our purpose of identifying the most central patents and companies within several periods. Indeed, the same approach can be used to compare within-period firms' centrality on the main paths. This is done in two steps. First we calculate the average PathC index obtained by the subset  $N_{jT}$  of patents belonging to company  $j$  at period  $T$ :

$$PathC_{jT} = \frac{1}{N_{jT}} \sum_{i=1}^{N_{jT}} \frac{SPNP_i}{\frac{1}{N_T} \sum_{i=1}^{N_T} SPNP_i} \quad (3.4)$$

Being based on the average of patents' centrality, his aggregating method is biased against large firms. The reason is straightforward. Large firms sample more patents from the population. Therefore, the larger they are, the more the centrality of their average patent will approximate the average SPNP for each grant year. Therefore, the aggregate PathC will approach to one for  $N_{jT}$  approaching  $N_T$ , shadowing firm's *real* centrality. Hence, to use the PathC as a measure of companies' centrality we need to correct for this bias by removing the firm's size effect. We do that by bootstrapping firms' average PathC. For each firm, we randomly sample  $N_{jT}$  patents out of the population of patents granted in the given period under observation and then calculate the average PathC of the random sample. We repeat this process for 1000 iterations. This returns a distribution of expected PathC for each firm, given its size and the topography of the network in each period under observation. This distribution provides the p-value for the empirically observed PathC. We can then assess whether the company is significantly more (or less) central than expected by its size and the structure of the system of technologies. We then assess the strength of firms' centralities (or the lack of it) by computing z-scores of the empirical PathC compared to the expected one.

$$z - PathC_{jT} = \frac{PathC_{jT}^{empirical} - \mu_{PathC_{jT}}^{random sample}}{\sigma_{PathC_{jT}}^{random sample}} \quad (3.5)$$

The z scores of the PathC index can be interpreted as a measure of technological influence of a particular company. A large z-PathC for a given firm means that, in the given period, the firm was significantly more influential than expected in terms of shaping the direction of the paths of technical improvements.

### 3.3.4 Measuring path-following, path-changing, ranking-following and ranking-changing propensity

We argued that the direction of technological change can be explained by two dimensions, the ranking of engineering problems and the stability of approaches to problem-solving. Therefore, players' contribution to shaping the direction of technological change can be described according to these two dimensions.

Depending on the extent to which innovative effort is directed to previously attractive engineering problems or to newly attractive ones, economic agents can be *ranking-followers* or *ranking-changers*. This is measured by the observed odds that patents granted to a given agent connect to a system of engineering problems (i.e. the component of the NMPs) that is not the highest ranked one. The *rank-changing index* (RCI) for agent  $j$  at period  $T$ , is calculated as follows.

$$RCI_{jT} = \frac{PC_{jkT}^{k=2} - PC_{jkT}^{k=1}}{PC_{jT}} \quad (3.6)$$

Where  $PC_{jt}$  is the number of patents granted to agent  $j$  in period  $T$ ,  $PC_{jkT}^{k=1}$  is the count of patents that connected to the first component of the NMPs for company  $j$ , in period  $T$ , and  $PC_{jkT}^{k=2}$  is the same for the second component. In other words, the RCI is simply the difference between the share of firm  $j$ 's patents that connect to the first and second largest components of the NMP in period  $T$ . Therefore, a large RCI indicates that the given agent is generating more inventive output in the second largest system of technologies than in the first, contributing to change the ranking of the system of engineering problems. Using shares allows taking into account how spread firm's inventive activities are across all the components of the NMP. Yet, similarly to the PathC index, we first need to correct for a bias against large firms that is intrinsic to the way we defined the RCI. Just by random chance, large firms can be expected to have a more broadly diversified set of inventive activities. This is because the more patents they randomly sample out of the population the more the set of inventive output is likely to be spread across components. Therefore, the larger  $PC_{jT}$ , the smaller the shares of output in the first and second component and, consequently, the smaller the RCI. Once again, we rely on bootstrapping techniques to clean for this size-induced effect. For each firm  $j$  in each period  $T$ , we randomly sample  $PC_{jT}$  items from the population of patents under the null hypothesis that each component has the same chance to be selected at each draw. We repeat this process for 1000 iterations and each time compute the RCI. We then compute p-values and adjusted z-scores (i.e. cantered on the mean of the random sample) for the empirically observed RCI. This exercise allows comparing the observed RCI with the one expected assuming that firms have no specific interest in any area of the system of engineering problems. Therefore, the z-score reveals how much a firm truly focus on the second component of the NMPs compared to the first one, given its size. Thus, it can be used to identify significant ranking changers and followers.

With respect to the approach to problem-solving agents (be they patents or firms) can be *path-followers* or *path-changers*. To quantify the extent to which an agent is following or changing the approach to problem solving we need to look at which prior-art she is building upon. We measure this by means of the *path-following index* (PFI). The PFI for patent  $i$  at time  $t$  is measured as follow.

$$PFI_{it} = \frac{1}{N_{S_{it}}} \sum_{j \in S_{it}} PathC_j \quad (3.7)$$

Where  $PathC_j$  is the average Path Centrality Index of the cited patent  $j$ .  $N_{S_{it}}$  is the size of the set  $S_{it}$  of patents cited by patent  $i$ . Therefore, the PFI is the average path-connectivity of  $i$ 's cited patents. If patent  $i$  connects on average to patents that are more central than their peers (i.e. with  $PathC > 1$ ),  $i$ 's PFI will be larger than one. Oppositely, a PFI smaller than one reveals that patent  $i$  connected to patents on average less central than their peers, namely to a relatively unexploited prior art. In particular, their low PathC indicates that they build on a short path of improvements and that their technical solutions had been largely unexploited, at the time the new patent connects to them. This is similar to the concept of *exploration* defined as introducing solutions that are *new-to-the-world* (March, 1991). The strength of path-following or path-changing behaviour at the patent level, can be measured by computing a p-value statistic. To do that, for each five-year period  $t$ , we estimate the cumulative distribution function of  $PFI_i$  and then we derive the PFI thresholds corresponding to the bottom five-percent of the distribution (for significant path-changers) and the top five percent (for significant path-followers).

Now we have a measure that quantifies the extent to which new patents connect to poorly or well-connected ones. Yet, citing poorly connected patents does not necessarily improve their



connectivity on the trajectory significantly. For that to happen, the citing patent must effectively set a new path to which many additional patents will connect in the future. This will eventually pull the previously poorly connected patents to the main paths. The *potential power-of-pull* index (POP) for patent  $i$  is computed as follows.

$$POP_i = \frac{R_i^{fwd}}{\frac{1}{n_t} \sum_{i=1}^{n_t} R_i^{fwd}} \quad (3.8)$$

Where  $R_i^{fwd}$  is the number of patents younger than patent  $i$  that reach him. This equals the number of patents that cite directly or indirectly patent  $i$ . The POP index measures how many forward (direct and indirect) connections a patent receives more than the average of his grant year cohort  $n_t$ , within the window of time under observation. We stress that this is a potential power of pull as the effective ability to pull unexploited prior art to the main paths depends on the long-term success of the new path. The POP only measures promising early success within a five-year period, as the number of direct and indirect forward connections is measured within the window of time under observation. However, patents with significantly larger POP receive more forward connections than their average peer does within the five-year period. This means that solutions disclosed in those patents are quickly attracts further inventive efforts, perhaps because long awaited. The p-value statistics for the POP can be calculated in the same vein as for the PFI, by first estimating its cumulative distribution and then extracting the threshold corresponding to the top five-percent. Analysing which patents have statistically larger POP and PCI allows identifying path-changing patents, as the citing patent is effectively changing the trajectories of problem-solving activity by pulling old patents that were previously unexploited into the NMPs. The same can be done at the firm level by calculating the average POP and PCI for each firms' set of patents. However, the aggregation of PFI and POP at the firm level faces the same size-induced bias discussed for PathC and RCI. Once again, this is solved by bootstrapping PFI and POP to remove the size effect, using the same procedure discussed above. Therefore, the t-statistics that we eventually use at the firm level are the following.

$$z - PFI_{jT} = \frac{PFI_{jT}^{empirical} - \mu_{PFI_{jT}}^{random sample}}{\sigma_{PFI_{jT}}^{random sample}} \quad (3.9)$$

$$z - POP_{jT} = \frac{POP_{jT}^{empirical} - \mu_{POP_{jT}}^{random sample}}{\sigma_{POP_{jT}}^{random sample}} \quad (3.10)$$

Where  $PFI_{jT}^{empirical}$  and  $POP_{jT}^{empirical}$  are the empirically observed averages of the values of PFI and POP for the set of patents belonging to firm  $j$  in period  $T$ . The larger the patent count for firm  $j$  in period  $T$ , the more the distribution of PFI and POP calculated for the random sample tends to approach normality. Therefore, firms with  $z - PFI \geq 2$  and  $POP \geq 2$ , are significantly more path-followers than expected by their size and the topology of the network in the period under observation. Similarly, levels of  $z - PFI \leq -2$  and  $POP \geq 2$  disclose significant path-changing firms.

### 3.4 Data

We conduct our analysis using data from the second version of the NBER Patent Citation Data File (Hall et al., 2001). This dataset contains a detailed classification of patents granted by the US Patent Office (USPTO) between 1975 and 2006, and includes information on patent citations. Records are classified by US patent class (*nclass*), International Patent Classification (IPC) codes, application year, grant year and the country where the assignee company is registered. Semiconductor technologies belong to the macro-category “electronics” of the US classification system. They are classified into five different subclasses. They are the followings:

- 257: Active solid-state devices (e.g. transistors, solid-state diodes)
- 438: Semiconductor device manufacturing: process
- 326: Electronic digital logic circuitry
- 505: Superconductor technology: apparatus, material, process
- 716: Design of semiconductor devices

The most interesting distinction is the one between classes 257 and 438. The former includes inventions related to particular semiconductor devices (transistors, solid-state diodes, integrated circuits, etc.) so it can be generally understood as the container of product innovations. Manufacturing process innovations are generally classified in class 438. These two classes together include about 86% of patents in our dataset in the period 1976-2006. Class 438 account for about 46% of semiconductor patents, whereas about 40% of them belong to class 257. Class 326 takes account of inventions associated to electronic circuits performing logic operations, which are those features that allow programmability of integrated circuits. It accounts for 7% of patents. Class 716 encompasses semiconductors design related inventions. 4.8% of patents are classified in this class. Finally, Class 505 comprises inventions related to the materials exhibiting superconductivity (of which semiconductors are made) and processes related to treating these materials. Only 1.6% of semiconductor patents belong to class 505. To construct the initial dataset of semiconductor patents we extracted from the complete NBER patent citations data file all patents belonging to one of the technological sub-classes listed above. Then we retrieved all citations for which both the citing and the cited patent belong to one of the five semiconductor-related classes. This returns 118361 patents and 779083 citations, covering the time span between 1976 (intended as the grant year of the patent) and 2006. This is what we refer to as the initial dataset of semiconductor patents.

#### 3.4.1 Descriptive statistics of the Network of Main Paths

As explained in Section 3.3.1, the initial dataset of semiconductor patents is further refined to single out only influent patents from the point of view of technology development. This is done by applying the Network of Main Paths approach. One has to imagine this procedure as opening a Russian nested doll (*Matryoshka*), in which at every layer we reduce the network further. As in a *Matryoshka* every reduced network is fully contained in the one before.

Table 2 compares network sizes over cumulated periods. For the sake of our dynamic analysis we decided to split the time span into 6 cohorts whose cutting year are 1980, 1985, 1990, 1995, 2000 and 2006. Following what we discussed in Section 3.3.1, in

Table 2, as in most of the rest of this work, we use cumulated time cohorts. That means that each time cohort builds on and includes the one before. We start from the initial dataset of 118361 patents and we remove the isolated ones (those who do not cite and are not cited by any other patent). The result, for the complete period 1976-2006, is a network of 114097 nodes.

This network contains several weak components<sup>11</sup>, which are by definition disconnected with each other. We extract the main component of the whole network and feed it into the main path algorithm. The NMPs itself has several components. To analyse changes in the ranking of the system of engineering problems, in the rest of the paper we focus on the two largest components of the NMPs. The main component of the NMPs for the period 1976-2006 is made of 3544 patents and 3562 citations. These are the most influential patents from the point of view of the paths of technical improvements in the semiconductor technology domain.

**Table 2: Comparison of networks' sizes**

|  | 1976-1980 | 1976-1985 | 1976-1990 | 1976-1995 | 1976-2000 | 1976-2006 |
|--|-----------|-----------|-----------|-----------|-----------|-----------|
| Whole network - number of patents                            | 2079      | 5631      | 12533     | 26853     | 54086     | 114097    |
| Whole network - number of citations                          | 2712      | 13310     | 40255     | 102957    | 272843    | 779076    |
| Main component -number of patents                            | 1703      | 5385      | 12348     | 26686     | 53874     | 113756    |
| Main component -number of citations                          | 2469      | 13164     | 40145     | 102864    | 272728    | 778890    |
| Network of Main Paths - number of patents                    | 1445      | 3490      | 6042      | 10107     | 15387     | 23428     |
| Network of Main Paths - number of citations                  | 1403      | 3291      | 5697      | 9489      | 14588     | 22077     |
| Network of Main Paths - Main Component - number of patents   | 694       | 1540      | 2678      | 2043      | 4557      | 3544      |
| Network of Main Paths - Main Component - number of citations | 756       | 1597      | 2734      | 2064      | 4617      | 3562      |

The NMPs is made of several components. Table 3 reports the number of patents belonging to each of the first five NMPs' components as a percentage of the total number of patents in the NMPs. If we look at the relative size of the main component one can immediately notice a drop from period 1976-1990 to 1976-1995. Until the end of the 1980s, the main component comprised at least 44% of the total patents in the NMPs. This percentage decreases significantly from the early 1990s onwards. Moreover, the same drop in size is observed at the aggregate level for the five largest components. In the first period of our sample, the first five components accounted for about 64% of the patents in NMPs. This percentage reduces by almost a half in the last period. This is a first interesting finding that suggests how the ranking of engineering problems increasingly changed from the beginning of the 1990s. We will analyse this in more details in the next section.

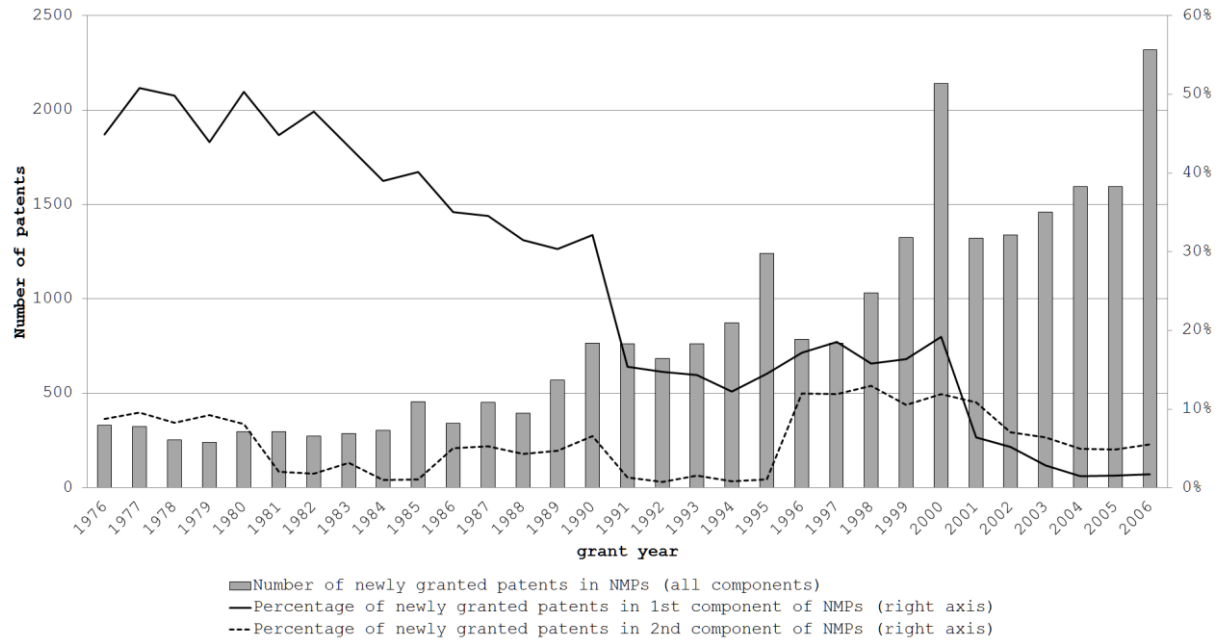
**Table 3: Relative size of the largest NMPs' component**

|           | First  | Second | Third | Fourth | Fifth | SUM    |
|-----------|--------|--------|-------|--------|-------|--------|
| 1976-1980 | 48.03% | 8.79%  | 3.32% | 2.01%  | 1.73% | 63.88% |
| 1976-1985 | 44.13% | 2.89%  | 2.24% | 1.60%  | 1.38% | 52.24% |
| 1976-1990 | 44.32% | 5.58%  | 1.26% | 1.22%  | 1.16% | 53.54% |
| 1976-1995 | 20.21% | 8.71%  | 6.57% | 6.33%  | 2.95% | 44.77% |
| 1976-2000 | 29.62% | 10.26% | 2.00% | 1.90%  | 1.37% | 45.14% |
| 1976-2006 | 15.13% | 11.79% | 3.55% | 2.01%  | 1.96% | 34.44% |

<sup>11</sup> In a directed network a weak component is defined as the subset of patents for which there exists a path from any node to anyone else once we symmetrize the dyadic relationship between nodes (i.e. we consider every link as bidirectional).

### 3.5 Findings

The indicators described in the previous section can now be used to analyse the direction of technological change, agents' contribution and the effectiveness of catching-up strategies. We present our findings in the next three sub-sections.



**Figure 12: Attractiveness of the two largest components of the NMPs**

#### 3.5.1 Direction of technological change

Figure 12 shows the evolution of the attractiveness index for the two largest components of the NMPs. The figure clearly shows that the largest component of the network of main paths dramatically loses attractiveness over time. In contrast, since 1996, the second largest component begins attracting more patents and overtakes the main one from 2001 onwards. This finding reveals a progressively strong change in the ranking of the system of engineering problems. It is interesting to report that the abstracts of the patents belonging to the second largest component of the NMPs in the 2000s, reveal a focus on displays and energy-saving technological solutions. This suggests that the second largest component of the NMPs is composed of technological areas more related to flat screens and portable devices rather than to desktop computers and laptops. This corresponds to a case in which a change in the technological trajectory, caused by a change in the functions of products that attract inventive efforts, disrupts the ranking of engineering problems. This is because the technical requirements of new portable devices pose different technological problems than designing and manufacturing PCs and laptops.

The ranking of the system of engineering problems therefore clearly changed over time. The graphical illustration of the main component of the NMPs and the visual analysis of its changing composition are reported in the Appendix A.3.1.

**Table 4: Patent persistence within the main component of the network of main paths**

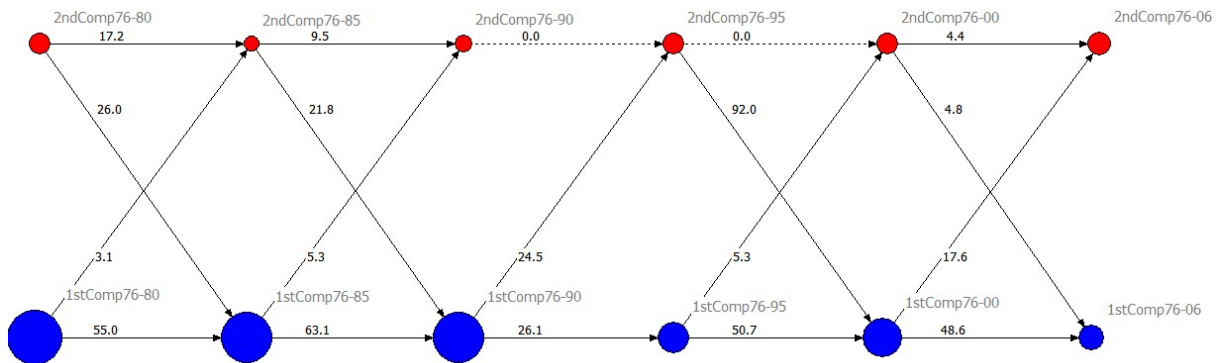
|           | 1976-1985 | 1976-1990 | 1976-1995 | 1976-2000 | 1976-2006 |
|-----------|-----------|-----------|-----------|-----------|-----------|
| 1976-1980 | 55.04%    | 62.97%    | 23.05%    | 62.39%    | 60.23%    |
| 1976-1985 |           | 63.12%    | 25.58%    | 66.04%    | 63.38%    |
| 1976-1990 |           |           | 26.14%    | 58.96%    | 53.06%    |
| 1976-1995 |           |           |           | 50.71%    | 35.93%    |
| 1976-2000 |           |           |           |           | 48.63%    |

Using the patent persistence index (PPI) discussed in Section 3.3.2, we can analyse stability in the approaches to engineering problem solving. Table 4 reports the PPI within the main component of the NMPs, decomposed by cohorts. In other words, the table shows the percentage of patents granted within the period reported on the rows, which are found in the main component of the NMPs computed for the period indicated in the column label. One can immediately notice that there has been a major disruption in the composition of the NMP in the first half of the 1990s. This means that patents granted in the period 1991-1995 were drawing on a previously not well-exploited prior-art. This reveals a change in the approach to problem solving. This happened despite the ranking of problems was stable in that period, as shown in Figure 12. Nevertheless, the disruption in the approach to problem solving has been lately offset by a subsequent counter-discontinuity that occurred in the second half of the 1990s. Indeed, patents granted in the period 1996-2000 reverted to exploit the old and paths of technical improvements established up to the end of the 1980s. In fact, about 63% of patents from the 1976-1980 cohort of the main component of NMPs are found in the main component in the period 1976-1990. Many of them are excluded from the main component in the period 1976-1995 (only 23% of them persisted in the main component) but reappear in the period 1976-1995, when the PPI returned to its 1976-1990 level. A very similar trend is observable for patents that belong to the main component in the period 1976-1985. Their level of persistence in the main component of the NMPs computed for periods 1976-1990 and 1976-1995 is almost identical to the PPI of patents belonging to the main component of the NMPs in period 1976-1980. Similarly, we observe that only 26% of patents from the main component computed in period 1976-1990 are part of the main component of NMPs76-95, whereas many more (about 59%) shows up in the last two periods. These findings clearly show that an important but temporary discontinuity in the approach to problem solving occurred in the first half of the 1990s.

So far, we have only looked at stability of problem solving approach in the main component of the NMPs. However, we know from the previous analysis that over time the main component has lost importance in favour of the second component. Therefore, to investigate stability of problem-solving approaches in the two main components, we analyse patent persistence within and across components with the help of Figure 13. Nodes on the first row represent the second component over time and those in the second row represent the main one. Node size is proportional to the relative size of each component in terms of number of patents (as reported it in Table 3). Links within and across components report PPI values.

Figure 13 reveals some interesting insights. First, the PPI within the second component is extremely low if existent at all. This means that the second component of the NMPs cannot be meaningfully interpreted as an established system of engineering problems alternative to the mainstream one. Yet it has an interesting role whose nature changes over time. In the first two periods, the second component hosts areas of research on semiconductor technologies that lately converge into the main paths of technical improvement that constitute the main component.

This means that, up to the end of the 1980s, the main paths of technical improvements partly built on prior art that was previously explored but still underdeveloped, whose usefulness was initially overlooked. In the two periods 1991-1995 and 2000-2006, the role of the second component changes. From source of novelty becomes a host for approaches to problem solving that gets temporarily disconnected from the main component. Indeed, in the period 1976-1995 it receives almost one fourth of the prior art contained in the main component in period 1976-1990. This set of prior art almost entirely converge back into the main component in the following period. Something similar happens between the last two periods, when almost eighteen percent of the patents found in the main component persist to the second one.

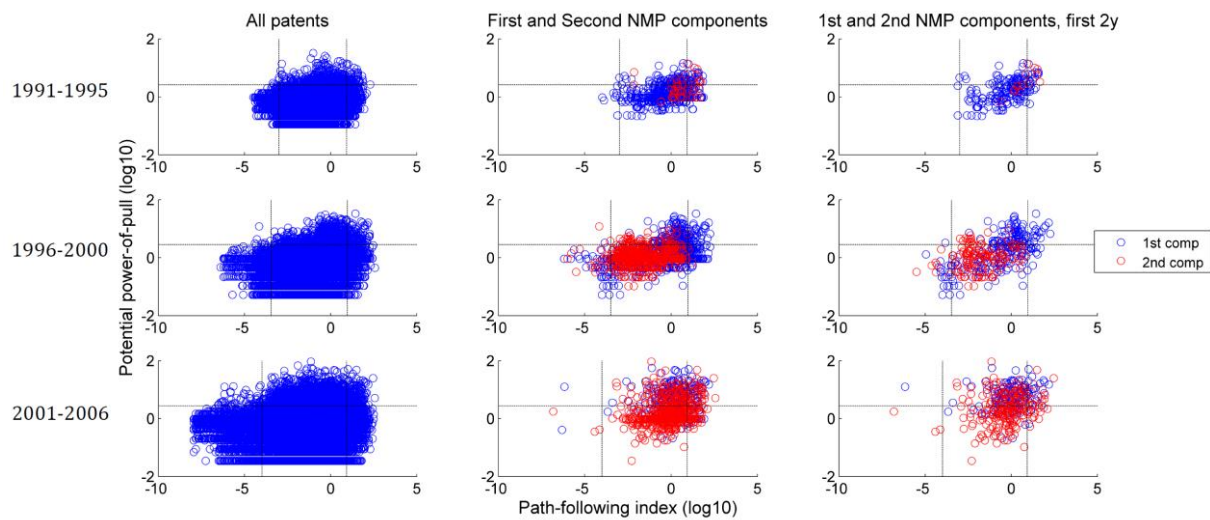


**Figure 13: Patent persistence within and between components**

We can derive two conclusions out of the analysis of stability of problem solving approaches. First, a major but temporary discontinuity in the paths of technical improvements occurred in the mid-1990s. Second, disagreement on which paths of technical improvements deserved more attention also arose in the first half of the 2000s. In both instances, the result was an increase in the breadth of search across useful prior-art, which resulted in the movement of a large set of patents from the main to the second component of the NMPs. The fact that those patents were still found in the NMPs, albeit in the second component, show that the prior-art that they represent was still perceived of some use by some players, although its importance decreased in relative terms. These conclusions highlight the heterogeneity of players' strategies and believes, which make technological change a complex dynamic. We analyse players' role and the effectiveness of their inventive strategies in the next section. This will also reveal who created those discontinuities. However, before to move to the firm-level, we can identify path-changers and path-followers at the patent level. We do that with the help of Figure 14, which reports patent position in the PFI-POP space. We focus on the last three five-year periods, 1991-1995, 1996-2000 and 2001-2006 as the previous analysis has shown that interesting changes in the direction of technical change have occurred in the semiconductors from the 1990s onwards. Panels in the first column, report the scatter plots of patents in the PFI-POP log-log space. They include all semiconductor patents granted in the given five-year period. The second column includes only patents located in the first two components of the NMPs granted in the same window of time. The third one includes only patents located in the first two components granted in the first two years of the five-year period. This is used as a robustness check. Even though the POP index is normalized by the average for each grant year, patents granted in the last three years might not have sufficient time to receive any forward connection and therefore the few that receive some might have a very high POP. This would add noise and might bias the results toward the youngest patents. Therefore, the plots reported in the third column are useful to assess whether our

results are strongly affected by patents granted in the last years of the five-year period. Clearly, they are not. In each panel, dashed lines mark the threshold for statistical significance. Patents located above the horizontal line and to the left of the first vertical one have a POP significantly higher than their peers in the five-year period and a significantly smaller PFI. Therefore, they are significant path-changers. Patents located above the POP threshold and to the right of the second vertical dashed line are significant path-followers. Panels in the second column show that nearly all the significant path-followers and most of the significant path-changers are located in the first two components of the NMPs, confirming the effectiveness of the NMP approach in identifying influential patents. Figure 14, also shows that significant path-changing patents are rare, much rarer than significant path-followers. This is not surprising as it is comparatively much more difficult to explore new paths *and be successful* (i.e. attracts more connections than expected), then successfully exploiting previously well-defined paths. There are many potential path-changers, but most of them do not have sufficient power of pull to effectively change the main paths of improvements. Also worthwhile noticing how, despite the previous analysis has shown that the engineering trajectories temporally changed in the first half of the 1990s, there are only two significant path-changing patents in this period. Whether they are alone responsible for the whole shift in the main paths of the network depends on how many patents they reached through indirect citations and of how many outgoing paths are generated by them. However, it is important to highlight how a shift in trajectories at the macro-level could also be the result of the accumulation of many small shocks. Namely of a large number of poorly successful patents that connects to the same small set of poorly connected patents.

Table 5, reports the list of path-changing patents. Reading these patent documents is very revealing about the effectiveness of our method to identify solutions that significantly change the approach to problem solving. Indeed, all of these patents describe *alternative* methods of solving *existing* problems, legitimizing the definition of path-changers. Two path-changing patents have been granted between 1991 and 1995. Both of them are located in the main component of the NMP. The one with lowest PFI is a Xerox 1992 patent on thin-film transistors, a technology that increases image stability and contrast in modern LCD displays. The patent describes an invention that uses a polycrystalline diamond film, which naturally exhibits light insulating properties. This allowed removing light shielding layers from thin film transistors, reducing device complexity as well as manufacturing time and costs. The second one describes a semiconductor device with several die stacked one on another within a single package. This ingenious solution allows larger computational power by eluding the problem of increasing density of integrated circuits (i.e. larger number of transistors per square centimetre of die area). The patent's assignee is single inventor, Mr. Peter J. C. Normington. This patent has been very successful in establishing a new path and has been cited 156 times.



**Figure 14: Identifying significant path-changer and path-follower patents**

For the second half of the 1990s, three patents can be identified as significant path-changers. The first one is a Bell Communications Research (BCR) patent that describes a solution to a high growth temperature problem related to complementary metal-oxide semiconductor (CMOS) processing. Previous solutions to temperature limitations had been proposed before the BCR patent, as acknowledge in its text. However, while effectively addressing temperature problems, previous approaches also caused an insufficient electrical conductivity of the heterostructure of  $\text{SiO}_2$  /Si substrates that made it insufficient for effective integrated circuit application, as explained in the BCR patent document. The BCR invention uses a different approach to address the high temperature limitations while providing the desired electrical sheet conductivity. The second path-changing patent in the 1996-2000 period is a Mitsubishi patent that describes a device having a multilayered metal interconnection structure with improved flatness. This solution, is also addressing the problem of generating higher speed and computational power by providing multiple layers of electrode interconnections. This solution set a path that will be eventually taken by organizations from latecomer countries, such as Taiwan Industrial Technology Research Institute (ITRI), foundries like Taiwan Semiconductor Manufacturing Company (TSMC), United Microelectronic Corporation (UMC) and Singapore-based Chartered Semiconductors, and Korean firms Samsung and LG Semiconductors. All of these firms have patents that cite Mitsubishi's technical solution. Successful US firms, such as Texas Instruments, Advanced Micro Devices (AMD) and Micron Technologies, also cites Mitsubishi's patent. In particular, Micron and AMD will eventually dominate the main two components of the NMPs in the period 2001-2006 in terms of centrality. This proves how much influence this new path has had in the process of technological change in the industry. The third path-changing patent in the second half of the 1990s has been granted to Bosch, for disclosing an invention that allows producing a hybrid semiconductor structure while tackling a main disadvantage of previously known methods in terms of reliability.



**Table 5: List of path-changing patents**

| period    | patent  | gyear | PFI    | POP    | #fwdcit | NMP<br>comp. | title   | assignee                     |
|-----------|---------|-------|--------|--------|---------|--------------|---|------------------------------|
| 1991-1995 | 5099296 | 1992  | 0.0007 | 4.571  | 15      | 1            | Thin film transistor  | Xerox                        |
| 1991-1995 | 5281852 | 1994  | 0.0027 | 4.016  | 156     | 1            | Semiconductor device including stacked die  | Peter J. C. Normington       |
| 1996-2000 | 5519235 | 1996  | 0.0001 | 3.286  | 45      | 1            | Polycrystalline ferroelectric capacitor heterostructure employing hybrid electrodes                       | Bell Communications Research |
| 1996-2000 | 5763954 | 1998  | 0      | 12.07  | 46      | 2            | Semiconductor device having multilayered metal interconnection structure and manufacturing method thereof | Mitsubishi                   |
| 1996-2000 | 5866951 | 1999  | 0.0003 | 3.7776 | 18      | 2            | Hybrid circuit with an electrically conductive adhesive   | Bosch                        |
| 2001-2006 | 6291319 | 2001  | 0      | 12.64  | 59      | 1            | Method for fabricating a semiconductor structure having a stable crystalline interface with silicon       | Motorola                     |

Finally, in the period 2001-2006, only one significant path-changing patent has been identified by our method. It is a process innovation developed by Motorola. The invention relates to a method for fabricating a semiconductor structure including a crystalline alkaline earth metal silicon nitrogen based interface. As explained in the patent document, this provides a more stable silicon surface, which is essential for subsequent epitaxial growth of single crystal thin films on silicon for numerous device applications. It is interesting to note that all of the path-changing patents identified by our method, except for one, belong to large incumbent corporations. This somehow constitute an argument in favour of the late Schumpeter's hypothesis that large firms are more conducive for radical innovations.

**Table 6: Top path-follower patents**

| period    | patent and assignee | gyear | PFI   | POP    | #fwdcit | PathC | NMP comp | title   |
|-----------|---------------------|-------|-------|--------|---------|-------|----------|---|
| 1991-1995 | 5061647 Motorola    | 1991  | 30.64 | 5.5745 | 49      | 228.1 | 2        | Forming a conductive layer over a substrate having a gate dielectric,...  |
| 1991-1995 | 5015598 Philips     | 1991  | 32.64 | 8.9875 | 69      | 222.8 | 2        | Metal-insulator-semiconductor   |
| 1991-1995 | 5168072 Texas Inst. | 1992  | 38.41 | 6.9655 | 162     | 199.4 | 1        | Method of fabricating an high-performance insulated-gate field-effect transistor  |
| 1991-1995 | 5079180 Texas Inst. | 1992  | 10.32 | 13.495 | 91      | 194.4 | 1        | Method of fabricating a raised source/drain transistor  |
| 1991-1995 | 5015595AMD          | 1991  | 22.57 | 10.011 | 51      | 132.1 | 2        | Method of making a high performance MOS device having both P- and N-LDD regions using single photoresist mask           |
| 1996-2000 | 5534447 UMC         | 1996  | 97.10 | 12.623 | 38      | 515.1 | 1        | Process for fabricating MOS LDD transistor with pocket implant  |
| 1996-2000 | 5538913 UMC         | 1996  | 150.0 | 16.170 | 76      | 495.6 | 1        | Process for fabricating MOS transistors having full-overlap lightly-doped drain structure                               |
| 1996-2000 | 5670401 Vanguard    | 1997  | 181.5 | 6.692  | 19      | 546.5 | 1        | Method for fabricating a deep submicron mosfet device using an in-situ polymer spacer to decrease device channel length |
| 1996-2000 | 5801083 Chartered   | 1998  | 94.13 | 10.282 | 149     | 540.4 | 1        | Use of polymer spacers for the fabrication of shallow trench isolation regions with rounded top corners                 |
| 1996-2000 | 5489543UMC          | 1996  | 27.48 | 33.333 | 54      | 449.4 | 1        | Method of forming a MOS device having a localized anti-punchthrough region  |
| 2001-2006 | 6190977 TI-Acer     | 2001  | 15.04 | 8.160  | 26      | 344.1 | 2        | Method for forming MOSFET with an elevated source/drain   |
| 2001-2006 | 6524920 AMD         | 2003  | 42.98 | 12.616 | 30      | 326.2 | 2        | Low temperature process for a transistor with elevated source and drain   |
| 2001-2006 | 6703648 AMD         | 2004  | 53.18 | 20.805 | 55      | 364.3 | 2        | Strained silicon PMOS having silicon germanium source/drain extensions and method for its fabrication                   |
| 2001-2006 | 6979855 Micron      | 2005  | 32.28 | 5.1814 | 13      | 490.1 | 1        | High-quality praseodymium gate dielectrics  |
| 2001-2006 | 6921702 Micron      | 2005  | 38.02 | 14.063 | 67      | 316.9 | 1        | Atomic layer deposited nanolaminates of HfO2/ZrO2 films as gate dielectrics   |

It is also insightful to identify significant path-follower patents. Since there are many of them, in Figure 14 we only report information for the most central ones (i.e. the top-5 of the PathC ranking). Two major insights emerge from the table. First, some of the path-following patents listed in the table address similar research areas, as proved by the frequent appearance of the terms ‘source/drain’ in the titles and the reference to MOS/MOSFET/PMOS/CMOS devices. These related patents are found in the first component of the NMPs in 1976-1995, and in 1976-2000 but move to the second in 1976-2006. They are part of the same long path of improvements that relates to a key transistor technology, metal-oxide-semiconductor (MOS). The metal-oxide-semiconductor field-effect transistor (MOSFET) provides switching function in microprocessors and memories to implement logic gates and data storage. They are the most widely used type of transistor in integrated circuits. This explains why they attracted such a constant inventive effort over time (MOS technology actually dates back to the 1960s). Particular attention has been devoted to the fabrication method, to make them cheaper and facilitate integration. For instance, the technique called complementary-metal-oxide-semiconductor (CMOS) consists in coupling two complementary MOSFETs into one high/low switch. This technique increased modularity, by allowing separating IC design from manufacturing (Adams et al., 2013). One major problem, as explained in patents 5670401 (Vanguard) and 6190977 (TI-ACER), is that the constant effort to miniaturize devices implies the need to solve a bunch of related problems. As stated in the TI-ACER patent, *“as the MOS transistors become narrower and thinner accompanying with shorter channels, the problems like the junction punchthrough, the leakage, and the contact resistance cause the reduction in the yield and the reliability of the semiconductor manufacturing processes”*. The preparation of an extremely shallow source/drain junction is one way to solve some of these problems, which has attracted the inventive effort behind the solutions disclosed in the three patents located in the second component of the NMPs in period 2001-2006. These findings show how the approach to solve key problems, which arise from a long-lasting trajectory of miniaturization efforts, has been following a well-established path over the last four decades. This path is traced by solutions that largely shared the same approach to problem solving. The table also provides an interesting insight into the dynamics of catching-up. The ranking of path-follower patents for the period 1996-2000 is dominated by latecomer firms, whose inventions address some of the MOSFETs related problems listed in the quote above. The diffusion of CMOS and MOSFETs opened entry opportunities in IC manufacturing for latecomers, as discussed in Adam et al. (2013) and Langlois and Steinmuller (1999). This implied that latecomer innovative effort focused on fabrication-related problems. These problems became major bottlenecks for further IC performance improvements, due to increased chip miniaturization. This ultimately explains the centrality of latecomer’s solutions in the NMPs. Three of these firms are from Taiwan. United Microelectronics Corporation (UMC) is the second largest Taiwanese semiconductor companies, and second largest foundries in the world by revenues (according to IC Insights), behind Taiwan Semiconductor Manufacturing Company (TSMC). Vanguard is the third Taiwanese IC foundry service provider. It started in 1994 as a DRAM subcontractor of TSMC. It then evolved into a pure-play foundry company. TSMC and the Industrial Technology Research Institute (ITRI) are its largest investors. The last patent to appear in the top five path-follower ranking belongs to Chartered Semiconductors. Chartered Semiconductors was a Singaporean company which, prior to its acquisition by GlobalFoundries in 2010 (a joint venture between AMD and Abu Dhabi’s Advanced Technology Investment Company), was the world’s third largest dedicated independent semiconductor foundry, after TSMC and UMC. Another Taiwanese firm hold a key path-following patent in period 2001-2006. This is ACER. The company started as Texas Instrument-Acer, a joint venture between the US incumbent Texas Instruments and the Taiwanese

new entrant ACER. The joint venture was founded in 1976. Twenty-one years later, in 1997, Acer acquired the notebook division of TI and became one of the world leaders in notebook manufacturing.

We have analysed the direction of technological change in the Semiconductor industry along two dimensions, the ranking of engineering problems and the stability of problem-solving approach. This allowed identifying a discontinuity in the engineering trajectories of technical improvements in the first half of the 1990s and at the beginning of the 2000s, and a change in the ranking of problems in the early 2000s. We have also identified path-changing and path-following patents, which greatly shed light on the nature of these dynamics. We can now move to the organizational level and investigate which firms shaped the direction of technological change and which one followed the course of the events. We only focus on the last three periods, given that the analysis presented in Section 3.5.1, identified them as interesting cases of change in the direction of technological progress in the Semiconductors.

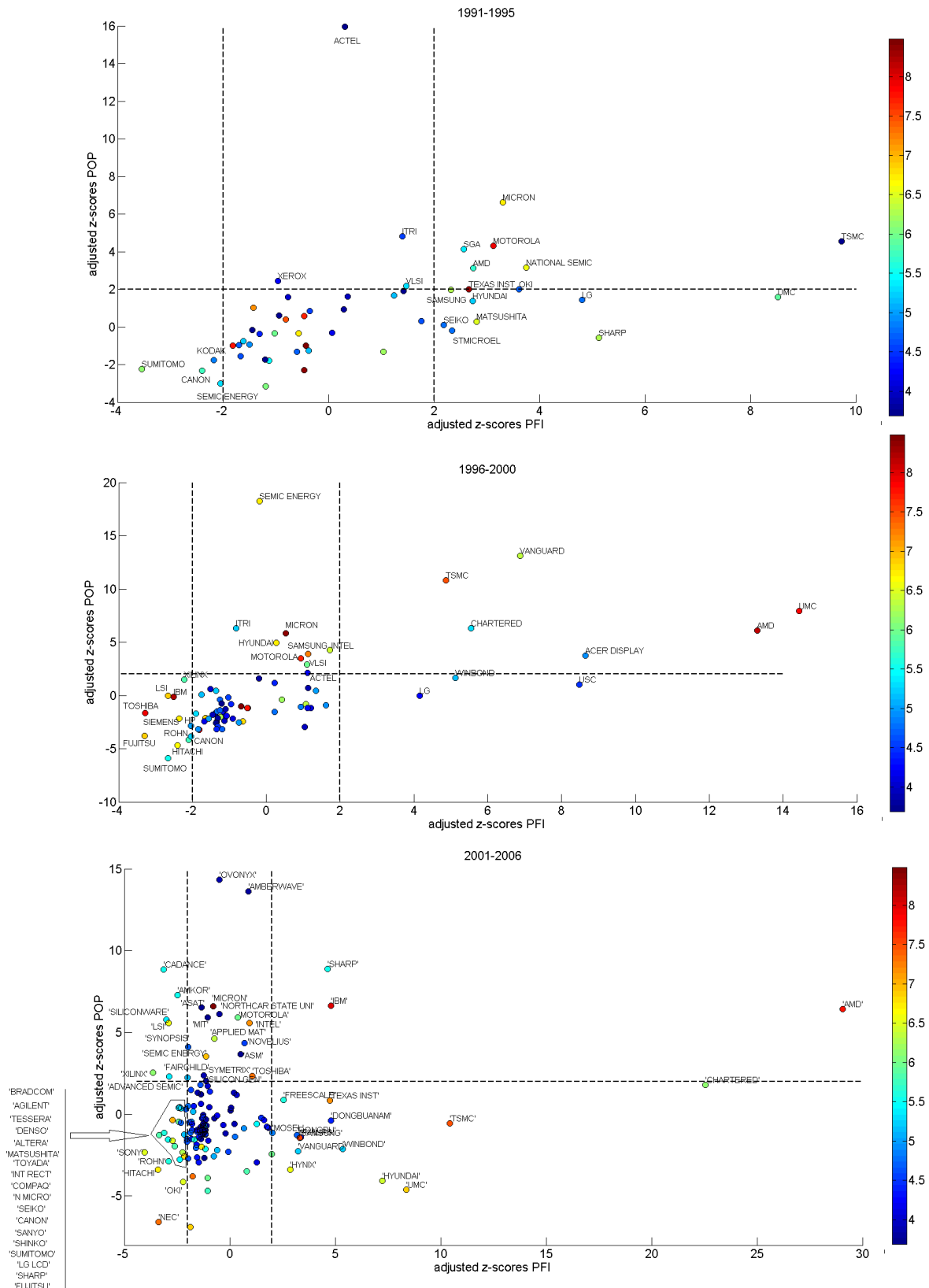
### 3.5.2 *Players' inventive strategies*

We have analysed the direction of technological change in the Semiconductor industry along two dimensions, the ranking of engineering problems and the stability of problem-solving approach. This allowed identifying a discontinuity in the engineering trajectories of technical improvements in the first half of the 1990s and at the beginning of the 2000s, and a change in the ranking of problems in the early 2000s. We have also identified path-changing and path-following patents, which greatly shed light on the nature of these dynamics. We can now move to the organizational level and investigate which firms shaped the direction of technological change and which one followed the course of the events. We only focus on the last three periods, given that the analysis presented in Section 3.5.1, identified them as interesting cases of change in the direction of technological progress in the Semiconductors.

Figure 15 shows firm position on the PFI-POP space. The value of the indices reported on the axes are the z-scores computed on the basis of randomized samples of equal firm size, as explained in Section 3.3.4. This allows removing any possible effect related to large disparities in the number of patents hold across firms. Marker colour is proportional to log-transformed patent count. Dashed lines indicates the thresholds of statistical significance, which lay two standard deviations above the mean (i.e.  $z=2$ ,  $p\text{-value} \cong 0.05$ ). For visual purposes, we only plot firms with a number of patents significantly greater than the geometric mean of firm patent count (i.e. two standard deviations of the log-transformed data). This allows focusing only on firms that dedicate a significant amount of inventive effort in Semiconductors. We report firm labels for firms located in areas of the space that fulfil statistical significance requirements. A number of key insights emerge from Figure 15.

From a methodological point of view, the comparison of observed PFI and POP with their random counterpart effectively removed scale effects. Large and small firms are scattered in the space without any particularly remarkable order. Size does not seem to affect neither PFI, nor POP. Consequently, firms' centrality (measured by PathC) is also independent from their size. We investigate if there is a relationship between the number of patents owned by a firm and the three indicators, PFI, POP and PathC in the Appendix A.3.2. The statistical analysis confirms that POP is independent from PFI and PFI, POP and PathC are independent from patent count. This is in itself an interesting insight. It shows that large players are as likely as small ones to have influential and high quality patents, and have same likelihood to be path-changers/followers. Moreover, being a path-follower does not necessarily ensures receiving more forward connections than

randomly expected, compared to being a path-changer. This reveals that, at the firm level, there is no preferential attachment of new patents to established paths. In turns, this means that, in principle, being a path-follower does not automatically make players more central.



**Figure 15: Identifying path-changing and path-following firms**

(marker color proportional to log transformed patent count)

In each of the three periods, the PFI-POP space is relatively more crowded in the area whose centrepiece has a PFI of -1 and POP of -2. This means that many semiconductor-focused firms seem to have a slight tendency toward being path-followers but in general have patents of a relatively poor influence (i.e. the average POP is smaller than expected by firm size). However, the strength of this tendency is not sufficient to pass the thresholds of statistical significance, showing that this pattern could possibly be just a random phenomenon. A few firms stand out of the crowd. Nevertheless, several of those that can be distinguished for unequivocally strong path-changing or path-following strategies, lack the strength to pull new paths into the network, or hold old ones in, alone. This is revealed by lower levels of POP than expected by firm size. Yet, a few insightful tendencies can be observed.

Firms that pass the requirements of statistical significance for both PFI and POP, are path-followers in the 1990s and mostly path-changers in the first half of the 2000s. This reveals that the temporary change in the engineering trajectories that happened in the first half of the 1990s, as discussed in Section 3.5.1, was not caused by a particular set of firms. Rather it was caused by idiosyncratic inventive outputs that explored new approaches. No firms were specifically trying to change the main trajectories in that period. In contrast, changes that happened in the early 2000s, were driven by specific strategies of a few firms that focused their inventive efforts to new paths and new problems.

Firms from latecomer countries (Taiwan, S.Korea and Singapore), have a strong preference to connect to established paths. This is especially clear for the world largest foundries, TSMC (Taiwan), UMC (Taiwan), and Chartered (Singapore, now called Global Foundries). They are always found in the right part of the space. They have PFI significantly larger than expected and in some cases even a POP above the threshold for statistical significance (TSMC in the 1990s, UMC and Chartered in 1996-2000). The group is joined by Vanguard, another Taiwanese foundry and Acer Display, at the end of the 1990s. Winbond, the largest brand name integrated circuit supplier in Taiwan, is also close to be a significant path-follower in the period 1996-2000. Korean firms, LG (in the 1990s), Hyundai Electronics and Samsung (both in 1991-1995 and 2001-2006), also have PFI much above the expected level, but the power-of-pull of their patents is not as strong as for the Taiwanese companies and Chartered. American microprocessor company Advanced Micro Devices (AMD), Intel's closest competitor, is also part of the path-following group, in each of the three periods. This contrasts with Intel's strategy, whose patents have a comparable power-of-pull but are much less focused on following the main paths of improvements than AMD's ones.

In the first half of the 2000s, a group of significant path-changers emerged. Some of them are relatively small players. Yet, some are noteworthy. Xilinx Semiconductors is famous for inventing the field programmable gate array (FPGA) and being one of the first semiconductor company to adopt the fabless business model. LSI Logic is another fabless focused that had an historical role in the industry. LSI helped create and nurture development of the application-specific integrated circuit (ASIC) industry. The company now focuses on designing semiconductors for storage, networking in data-centers, mobile networks and client computing. Both Xilinx and LSI Logic had a strong preference for being path-changers in the second half of the 1990s. However, only in the early 2000s they got a sufficient power-of-pull to become significant path-changers. The presence of these two important fabless, that are known to be central players in the industry, suggests that our method is indeed able to identify path-changers. Fairchild, Intel's mother, is also part of the path-changing group. Fairchild main line of business is now focused on high-performance power ICs and power discrete devices. The combination of path-changing companies working on power, mobile networks and networking, in the 2000s, reflect the shift-

ing target of applied research on semiconductor devices from desktop PCs to laptops, tablets, smartphones and cloud computing.

There is also a crowd of potential path-changers in the first half of the 2000s, which have very strong preference to connect to unexploited paths, but lack sufficient power-of-pull. Among them, some consistently show a path-changing strategy in more than one period. These are mostly Japanese incumbents such as Fujitsu, Hitachi, Canon, Rohm Semiconductor and Sumitomo.

As we broadly discussed in Section 3.2, the stability of approaches to problem solving, here represented by the tendency to be path-followers, is just one of the two crucial dimension of the direction of technological change. The ranking of engineering problems is also insightful about the characteristics of technical progress. We have seen in Section 3.5.1, that a change in the ranking has occurred in the 2000s. The RCI, allows identifying which players reported more inventive output in the second component of the NMP than in the first, compared to what could be expected by random chance. Figure 16 reports the RCI ranking for firms in the last three periods. We only show firms whose RCI is statistically significant at the 95%. Besides being path-followers, firms from latecomer countries were ranking followers. This is the case for UMC, TSMC, Chartered, Samsung and Vanguard. In the 1990s, large American and Japanese incumbents were those with relatively more inventive output in in the second component of the NMP, effectively trying to revert the ranking of the systems of engineering problems. This pattern changed in the 2000s. None of the latecomers is found in the list of significant ranking followers in the last period. In contrast, TSMC, together with the Korean Electronics and Telecommunications Research Institute (ETRI), Hann Star Display and LG LCD, were part of the group of ranking-changers. The latter two are specialized on display technologies. The same holds for another ranking changer, Semiconductor Energy Laboratories (SEL). SEL's president and majority shareholder, Shunpei Yamazaki, used to be the most prolific inventor according to the number of USPTO patents hold, until 2008. The presence of these firms confirms what emerged from the analysis of path-changers, namely that the second component of the NMPs in the first half of the 1990s, is characterized by the emergence of a new set of engineering problems. It is worth noting that the most ranking-changing player in the second half of the 1990s was Tessera Technologies. This US company is focused on miniaturization technologies that then aggressively license out to big players. Indeed Tessera is known as one of the largest patent trolls in the industry (Business Insider, 2012).

### 3.5.3 *Effectiveness of players' strategies*

Having analysed players strategies we can now assess their effectiveness in terms of earning centrality on the paths of technological improvements. Figure 17 reports the rankings of the z-scores of PathC at the firm level. We only show firms whose z-score is statistically significant (i.e.  $z \geq 2$ ). As we discussed in the previous section, our findings show that there is no relationship between PFI and POP. Therefore, in principle, it is possible to achieve high centrality even with a path-changing strategy, as long as the path-changing patents attract a considerable number of forward connections. However, in practice, it seems that path-following strategies are much more likely to earn centrality on the main paths. Indeed, only two of the most central firms listed in Figure 17, have a PFI lower than zero. These are Micron (PFI=-0.77) and Applied Materials (PFI=-0.6) in the last period. The cautious path-and ranking-following strategy of Taiwanese and Singaporean foundries has been very successful in terms of gaining centrality in the network, also due to their high quality patents (large POP). Since the early 1990s, TSMC and UMC have

been very central, and they maintained a position in the top ranking in the two subsequent periods. Chartered, Vanguard, Winbond and Taiwan Industrial Technology Research Institute also have been very central in one or two periods. This confirms the remarkable ability shown by these firms to quickly catch-up in terms of technological capabilities. Korean firms Samsung and LG also followed similarly cautious strategies, but they have been somewhat less successful in terms of centrality.

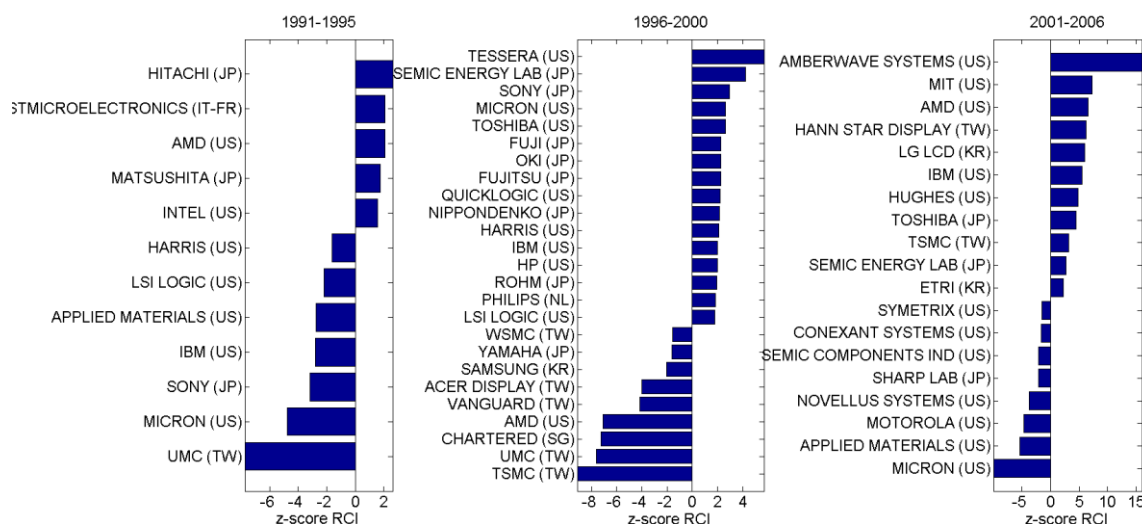


Figure 16: Ranking-changing and ranking-following firms

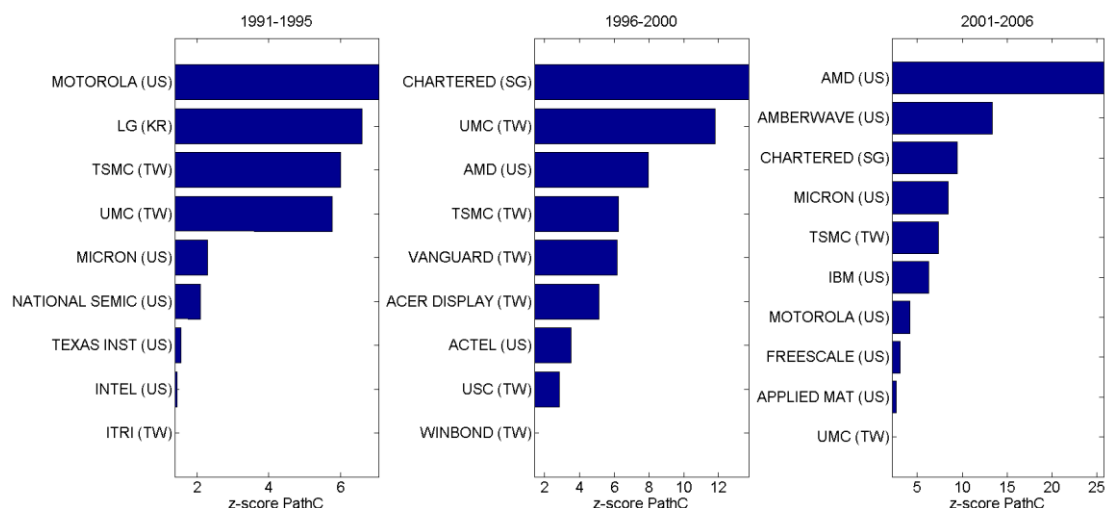


Figure 17: Most central players

### 3.6 Conclusions

In this work, we developed a theoretical and methodological framework to analyse two key dimensions of the direction of technological change, the ranking of engineering problems and the stability of approaches to problem solving. We applied this framework to the semiconductor industry, giving particular attention to the analysis of catching up strategies followed by latecomer firms. We showed that, a major change in the approach to problem solving has occurred in the



first half of the 1990s. Two patents were identified as significant path-changers. They focused on designing semiconductor structures made of polycrystalline diamond thin-films, and on vertical stacking dies on a single chip to increase computational power without increasing horizontal density of elements on microchips. We also identified a change in engineering trajectories, coupled with a change in the ranking of problems, which occurred in the first half of the 2000s. This was caused by advances in the miniaturization of semiconductor devices allowed by the diffusion of MOSFET technologies. This technology is the result of a long and well-established path of improvements, which could be referred to as a sort of backbone of the system of semiconductor technologies. However, the market success of new applications for semiconductor devices (like smart portable devices and LCD televisions), changed the connections among engineering problems, and created new challenges that attracted the relative majority of inventive efforts from the beginning of the 2000s.

Our findings show that, in the 1990s and first half of the 2000s, firms from latecomer countries (Taiwan, Korea and Singapore), especially Taiwanese and Singaporean foundries, were significantly more path and ranking followers than expected by their size and time of entry. However, some of them, notably those engaged in LCD technology, were significant ranking changers in the early 2000s. This path and ranking following strategy has been instrumental in granting these firms centrality on the main engineering trajectories. Indeed, patents granted to Taiwanese, Singaporean and, to a smaller extent, Korean firms, were of a relatively high quality, attracting significantly more forward connections than expected by their size and time of entry. This shows how these firms were able to quickly develop the necessary technical capabilities to contribute high-quality inventive solutions that further strengthened the main trajectories of engineering improvements. Moreover, the fact that some latecomers are ranking changers in 2000s suggests that the path-following strategy adopted in the decade before, which ensured them strong centrality on main paths of improvements, strengthen their technological capabilities. It provided experience and knowledge about the direction of technological change in the industry, which allowed them being actively changer of the ranking of engineering problems in the early 2000s. These findings are in line with the anecdotal evidence on catching-up strategies in the semiconductor industry, as shown by Mathews and Cho (1999), Chang et al. (1994) and Cho et al. (1998).

It is important to highlight that our definition of firm's strategies with respect to the paths of technical improvements loses one category compared to Lee and Lim's (2001) threefold classification of catching-up strategies. The authors classified latecomer firms into path-following, path-creating and path-skipper. The empirical identification of path-skipping strategies requires defining and tracing the whole life cycle of a given technology. In particular, in our current framework, we do not make any statement about the future success or failure of path-changers and the new trajectories of improvements that they contributed. However, this framework can be suitably modified for identifying technology domains and trace their life cycle. This is done in the next chapter. In the third chapter, we will instead analyse how changes in the main trajectories of engineering improvements affect firms' survival and how diversification can provide an evolutionary advantage over competitors in uncertain technological environments. Finally, constructing an index able to reveal statistically significant path and ranking changer has to be considered as a first step of a potentially very rich research agenda. It becomes possible, for instance, to study whether a common profile for significantly more explorative firms can be described. Our analysis has already showed that, at least for the semiconductor industry, size, measured by patent count, is not a good predictor of path and ranking-changing behaviour. Alternative explanations might point to particular market or technological conditions that push

these firms to be more risk-takers. Another hypothesis could be that path and ranking-changers have a more diversified research workforce that allows them to pursue several paths. Our indices make it possible to perform an econometric analysis with the goal of revealing predictors of firms' path and ranking changing behaviour and, consequently contributing to unveil the future direction of technological change.

# 4 LIFE CYCLE OF TECHNOLOGY DOMAINS AND COMPARATIVE TECHNOLOGICAL ADVANTAGE

Catching-up, leapfrogging and falling behind in terms of productivity in high-tech industries crucially depends on firms' ability to keep pace with technological change. In fast changing industries, today's comparative technological advantage does not guarantee tomorrow's success. Firms' innovation prospects depend on their learning paths as the set of central engineering problems and the approaches followed to seek their solutions change with time. This highlights the importance of studying the relationship between life cycle of technology domains and comparative advantage patterns of new and incumbent innovators. We understand domains as areas of applied research that share common engineering problems and follow similar approaches to problem-solving. We theoretically define the life-cycle of technology domains and its relation with product and industry life-cycles. We empirically identify technology domains and trace their life-cycle by mean of patent citation analysis. The methodology is based on the analysis of the age of the different domains and of the characteristics of their technological trajectories. We then investigate comparative technological advantage patterns of new and incumbent innovators. Our findings prove that new innovators have a comparative advantage in emerging and young domains. We also show that, up to the end of the 1990s, firms from latecomer countries specialized mainly in areas at the later stages of their life cycles.

## 4.1 Introduction

In this paper, we analyse the relationship between the life cycle of technology domains and the revealed technological capabilities of incumbent and new innovators. The striking example of sustained fast economic growth and enormous structural transformation that several countries like the Asian Tigers (Hong Kong, Taiwan, South Korea and Singapore) and some of the BRICS (Brazil, Russia, India, China and South Africa) have provided in the last half-century, have been explained by a variety of points of view. A widely accepted explanation points to the role played by technology and knowledge upgrading as engines of economic growth and sources of international competitiveness (Kim and Nelson, 2000). The development of internal technological skills and the access to foreign technology is the key factor behind the process of catching-up (Fagerberg and Godinho, 2005; Hobday, 2000; Perez and Soete 1988; Verspagen, 1991; Abramovitz, 1994). Technology is in continuous evolution and the direction and speed of technical change, by creating and replacing capabilities at different paces, determine the availability of entry and catch-up opportunities (Lee, 2013; Lee and Lim, 2001 and Dosi, 1982) and changes in industry structure (Breschi et al. 2000; Malerba and Orsenigo, 1997; Schmookler, 1962). The life cycle of technology domains is therefore a determining factor of the evolution of products and industries and of the fate of catching-up endeavours.

Although technology, product and industry lifecycles are conceptually and dynamically intertwined, confusing the three levels of analysis generates conflicting predictions on the specialization patterns of new entrants. Industry life cycle theory (Klepper, 1997, 1996; Afuah and Utterback, 1997; Jovanovic and MacDonald, 1993; Suarez and Utterback 1993; Utterback and Abernathy, 1975) predicts higher entrance to occur in the earlier stages of the life cycle. This is when there are plenty of technological opportunities and a dominant design has yet to emerge. Consequently, the entry barriers are weaker due to the lack of cumulative technical and market knowledge advantage. Innovation management literature has also extensively analysed specialization of new entrants with respect to industry and product life cycles. However, the latter is even more specific than industry life cycle theory in predicting the type of technologies that are instrumental for successful entrance. Christensen disruptive technologies are the favourite competitive battlefield of new innovators (Christensen et al., 1998; Christensen, 1997). There are two main conceptual puzzles in these branches of literature. First, these theories focus exclusively on entrance from advance countries. Second, the theoretical framework does not clearly distinguish at which level between industries, products and technologies the mechanisms behind the life cycle operates. The literature provides two alternative theoretical approaches that focus on global competition: international product life cycle theory and catching-up. The international economics literature on product life cycle (PLC), sparked by the seminal contribution of Vernon, predicts that latecomers are more likely to specialize in obsolete technologies that are progressively abounded by leader countries and whose production moves to developing countries to exploit their comparative advantage based on cheap labour (Vernon, 1966). Recent findings in this strand of literature follows Vernon's framework (Bergek et al., 2013; Karniouchina et al., 2013). Vernon's theory has raised some criticisms, which focused mostly on the fact that today's production is characterized by fragmented value chains, and modular technologies and can therefore happen in more places simultaneously. Catching-up and technology regimes literature emphasizes how innovative entrance depends on changes happening at the technology level, as the introduction of new technologies or radical change in existing ones create higher technological opportunities which, *ceteris paribus*, tend to favour the entry of new innovators (Lee, 2013; Lee and Lim, 2001; Breschi et al. 2000)

A unifying framework that provides a systemic perspective relating industries, products and technologies is provided by Murmann and Frenken (2006). Industries can be seen as collection of vertically and horizontally related products which themselves are made of several components whose design and manufacturing require distinct technologies. Industry life cycle therefore depends on the life cycle of the underlying set of products. There is a wide agreement in the literature that a key factor that shapes product life cycle is the emergence of a dominant design after a phase of fluidity that involves searching several possible design paths (Afuah and Utterback, 1997; Suarez and Utterback, 1993; Anderson and Tushman, 1990; Utterback and Abernathy, 1995). Yet products do not necessarily offer the best resolution to study search across the design space and the emergence of orthodoxy in the design approach. Products are systems of components and sub-components whose development follows own technological trajectories. Therefore, the life cycle of technology domains clearly affects product and industry life cycles. A micro-founded analysis of entrance and catching-up must necessarily focus on studying change at the technology level, as it is at this layer that learning happens. The goal of this paper is to improve our understanding of the relationship between technology domain life cycle and specialization patterns of new innovators, in particular in the context of technological catching-up. Following Dosi's definition of technological trajectories, we conceptualize technology evolution as the process of solving engineering problems (Dosi, 1982). This involves searching for solutions following different approaches. We argue that the emergence of an accepted approach to problem solving and the stability of the set of problems is the technology domain level analogy to the rise of a dominant design at the product-level. We define technology domains as areas of research that define a set of common technological problems that are tackled applying similar mindsets and toolboxes.

Despite the variety of theoretical contributions to the literature of technology life cycle and dominant-design, few attempts have been made to empirically trace the evolution of technology. Perhaps this explains the contradictory predictions found in the literature on the specialization patterns of new entrants, specifically with respect to the direction of technology evolution. The few notable contribution to the study of life cycles at the technology level are the work of Jaffe and Trajtenberg (2002) and Lee (2013). Jaffe and Trajtenberg (2002) analysed the average time lag between cited and citing patents. They found that, on average, the number of citations to a given patent rapidly increases up to 3-4 years after it has been granted. It then relentlessly decreases. Lee (2013) argues that the citation leg trend is a good proxy of the technology life cycle as it reveal for how long the piece of technical information represented by a given patent keeps being a useful source of knowledge for improvement of technology. We argue that to link the technology and the industry level it is necessary to analyse the life cycle of the system of technologies within an industry, rather than focusing on single sub-classes. Furthermore, looking only at the citation lag provides a measure of the speed of change but do not provide a picture of the scope and direction of change. In this paper we contribute a method to identify domains within a technology class and analyse their life cycle. We can trace the stage of their evolution by looking at changes in the attractiveness of the engineering problems pertaining the given domain and the stability of the approaches followed to tackle them. Our method is based on a dynamic analysis of complex patent citation networks. We focus on the semiconductor industry as a case study as it provides a particularly suitable ground for testing such relationship. Industry leadership has changed over time, because of different waves of successful latecomer entrance. The industry is characterized by a persistently evolving knowledge base, increasing global competition and short business cycles (Brown and Linden, 2009). Furthermore, given the focus of this paper, it is particularly interesting to notice that the technology life cycle of semiconductors

is considerably shorter than other industries, as shown by Lee (2013). This has been proposed by Lee as a key explanation of the success of catching-up efforts due to the speed of knowledge replacement. Therefore, it is crucial to understand in which semiconductor technology domains new entrants specialize, determine the stage of their life cycle and assess whether latecomers specialization patterns progressively upgrade to emerging domains. In particular, we answer the following research questions: (i) In which life-cycle stages new innovators have a comparative technological advantage over incumbents? (ii) Are there significant differences in the revealed technological advantage of new innovators from different countries?

We identify domains and trace their evolution by analysing patent citation networks. Patent are understood as proofs that an innovative solution to a selected engineering problem has been found. Citations highlight the approach followed to tackle the problem by identifying the prior-art the design process has built upon. We use data from the second version of the NBER patent citation database (Hall et al., 2001) which covers the window of time between 1976 and 2006. To reduce noise in the data coming from the highly skewed distribution of patents' technical and economic value (Gambardella et al., 2008; Hall et al., 2005; Reitzig, 2003), first we identify the set of the most influential patents from the point of view of the development of the main technological trajectories within the semiconductor technology class. For this purpose we use the main path approach originally developed by Hummon and Doreian (1989) and subsequently refined and applied in recent work by Verspagen (2007), Fontana et al., (2009), Martinelli (2008; 2009), Bekkers and Martinelli (2010). Within this set of patents, we identify several interrelated technology domains using a community detection method proposed by Newman (2004). Then we develop a methodology to describe the life cycle stages of these domains according to the attractiveness of their engineering problems and the stability of the approaches followed to seek the solution. The basic intuition is that the centrality of the problems pertaining a given domain decreases over time, while the stability of the approaches to problem-solving increases.

The structure of the paper is as follows. First, we present a short overview on the technology and industrial dynamics of the global semiconductor industry (Section 4.2), to make the reader familiar with the background of this study. Then we introduce the theoretical framework that we followed to define technology life cycle (Section 4.3) and the necessary methodological steps to identify technology domains and infer the stage of their life cycles (Section 4.4). Finally, we present the results that answer the two research questions (Section 4.5).

## 4.2 Technology and Industrial Dynamics of the Global Semiconductor Industry

We focus our study on the semiconductor industry. Semiconductors are the best example of a high-tech industry in which catching-up, and possibly leapfrogging, by former laggard countries like Taiwan, Korea, and Singapore, prominently occurred. Technological change played a crucial role in determining the availability of entry opportunity. From the beginning of the 1980s onwards, the increased modularity of semiconductor design and manufacturing technology fragmented the value chain fostering specialization and jeopardizing production. New firms could now enter the industry at different stages of the production process. As argued by Adams et al. (2013), *"the increased adoption of Complementary Metal Oxide Semiconductor (CMOS) production processes weakened the interdependence of product design and manufacturing. [...] With the creation of standardized interfaces between components and Electronic Design Automation (EDA) tools a modular system developed. [...] The interdependence between product design and manufacturing was weakened in many product segments in semiconductors and specialist firms were able to enter the industry at both the design and the manufacturing stages"* (Adams et al., 2013, p.287). Fur-

thermore, on the product side the development of Application Specific Integrated Circuits (ASICs) and of systems on a chip (SoC), which squeezed all components of an electronic system into a single chip, allowed application customization. This fragmented the market and created entry and survival opportunities for small firms (Fontana and Malerba, 2010; Ernst, 2005; Linden and Somaya, 2003). Further miniaturization was also made possible by technological advancements in industrial lasers and lithography, which allowed exploring new innovative solutions. This shows how changes in technology strongly affect industrial dynamics and provides a first argument in favour of analysing the life cycles of semiconductor technology domains. A second one is provided by the short product life cycle that characterizes the industry (Brown and Linden, 2009). Periods of sustained revenues growth are constantly followed by periods of decline. This is also explained in a report by Integrated Circuit Engineering Corporation (ICE, 1996):

*"[In the] long term, the sustained profitability of the semiconductor manufacturers depends on each company's ability to maintain high enough profit margins on the devices it produces to allow sufficient capital outlays for future generations of devices. From year to year, the health of the semiconductor industry as a whole is indicated by its characteristic "boom" and "bust" periods, known as the silicon cycle. Since 1978, there have been four growth cycles in which sales grew an average of 30 percent per year. Following each growth cycle, the industry experiences a one to two year period when sales growth averaged slightly under 4 percent."* (ICE, 1996)

The cyclicity of the semiconductor industry at the business level provides a strong reason for studying its life cycle at the technological level as the introduction of new products depends on improving existing technologies and generating new ones. In this respect Lee (2013), by computing the citation lag between cited and citing patents, showed that semiconductor technologies have the second shortest life cycle length among those considered, with an average cycle time of 6.07 years before the underlying technical knowledge becomes outdated.

From the technological point of view there are a number of indicators that show as the semiconductor industry as a whole moved from an emerging phase in the second-half of the 1970s to maturity in the beginning of the 2000s. The empirical regularities of industry and product life cycles have been sketched by the work of Klepper (1996; 1997). The author summarizes them as follows: *"While distinguishing stages is somewhat arbitrary, the essence of the PLC is that initially the market grows rapidly, many firms enter, and product innovation is fundamental, and then as the industry evolves output growth slows, entry declines, the number of producers undergoes a shakeout, product innovation becomes less significant, and process innovation rises."* (Klepper, 1997, p.149). The shift in importance from product to process innovation is also at the hearth of models of industry evolution that followed the seminal contribution by Utterback and Abernathy (1975). The symptoms of a long-run maturity of the semiconductor industry are evident from Figure 18. Panel A reports the trend of the relative number of technologically influent patents granted by the USPTO from 1976 until 2006 in each of the five US sub-classes of the Semiconductor technology class (i.e. 438-process, 257-product, 326-materials, 505-programmability, 716-design)<sup>12</sup>. Product and process innovations were equally important in the end of the 1970s,

---

<sup>12</sup> It is important to note that the percentage of patents by technological class shown in Panel A of Figure 18 is not calculated based on all patents granted by the USPTO and classified in one of the semiconductor subclasses. Rather, we refer to the percentage by class with respect to a subset of technologically influent patents identified through the Main Path Approach, which we introduce in Section 4.4.1. The same holds for Panel B, where by new innovators we

when microprocessors had been recently introduced (Intel commercialized its first microprocessor in 1971), then process innovation strongly took over. Another sign of maturity, in line with life cycle theory is provided by the trend in technological concentration shown in Panel B. Industry maturity is usually associated with a decreasing number of new entrants (new innovators in our case) and an increasing concentration index. Panel B shows somehow contradictory evidence of that. The trend of innovative entrance appears to be quite cyclical, with two peaks reached in the second half of the 1980s and the 1990s. The number of incumbents, on the other hand, increases constantly (although at a decreasing pace) up to the end of the 1990s. This clearly points to the fact that some of the new innovators managed to successfully establish themselves in the industry. Consequently, the number of incumbents increased over time. Yet, in the first half of the 2000s, both the number of new innovators and incumbents decreases strongly. Consequently, the concentration index (we use the well-known Herfindal-Hirshman Index –HHI) explodes in the beginning of the 2000s. This reveals that the share of technologically influential patents had increasingly concentrated in the hands of a few firms. Therefore, at the technological level, the semiconductor industry is undergoing what is commonly defined as a shakeout in the 2000s, in line with the hypothesis of industry maturity. This is explained by life cycle theory as the result of the emergence of a dominant product design. There is anecdotic knowledge that semiconductor devices, such as microprocessors and memories, have indeed achieved great standardization and a dominant design focusing on miniaturizing their components has emerged (Epicoco, 2013). However, due to the multiproduct nature of the industry and the importance of customized application-specific semiconductors (ASICs), the theoretical concept of a dominant design, as defined in the literature, lacks meaning when broadly applied at the industry level. Yet, in Chapter 3, we have shown how a mainstream *approach* to design and a dominant system *of engineering problems* have guided innovative effort in the industry up to the end of the 1990s, as shown by the stability of the main technological trajectories<sup>13</sup> and the ranking of problems. Then, a new set of technical problems, mainly related to screen and MOSFET technologies arose, pushing designers to explore different approaches.

---

refer to firms that hold at least one patent in the subset of technologically influent ones for the first time and by incumbent innovators we mean firms that had at least one patent in that subset in the period(s) before.

<sup>13</sup> In the previous chapter, we show how the main approaches to engineering problem solving, i.e. the technological trajectories, actually got temporarily perturbed in the mid-1990s, when new players from S.Korea, Taiwan and Singapore became prominent innovators.



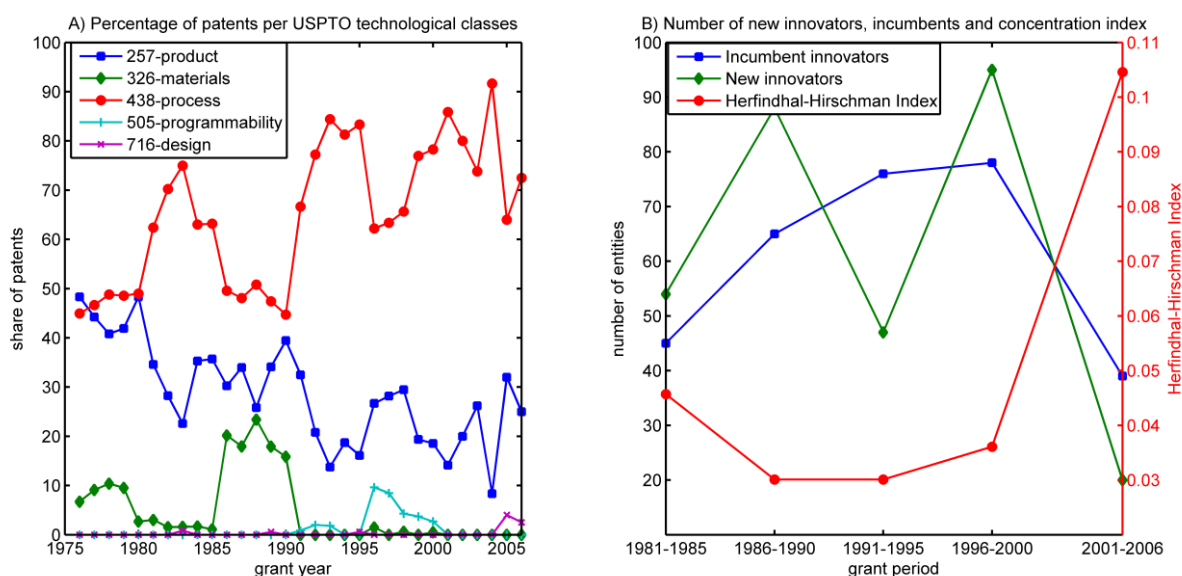


Figure 18: Industry technology dynamics

Given the record of successful catching-up efforts in the industry, it is interesting to brake the trend of new innovators of Figure 18 B at the country level. Figure 19 shows the share of new innovators by geographical origin. As we can see, innovative entrance is in accordance with the historical knowledge of the evolution of the global semiconductor industry, as described by Langlois and Steinmueller (1999). The share of US new innovators decreased over time up to the end of the 1990s, in favour of a larger entrance in the technological area by firms from Taiwan, Korea and Singapore, which account for about 20 per cent of all new innovators in the 1990s. In contrast, the share of new innovators from Japan is rather constant across our sample. Finally, it is interesting to note that, despite European firms becoming quite marginal players in the global semiconductor industry, they seem to be able to still play a significant role at the technological level, at least in terms of innovative entrance.

The brief overview of the industry trends reported in this section has shown how semiconductor firms are exposed to strong and shortening business cycles and how the industry seems to have entered in a phase of maturity in which only a few players are able to be successful in terms of generating influent technological innovations. These two effects reinforce the importance to clearly define life cycles at the level of technology domains and analyse specialization patterns of incumbent and new innovators. In the next section, we elaborate on the theoretical framework that guides us in this analysis.

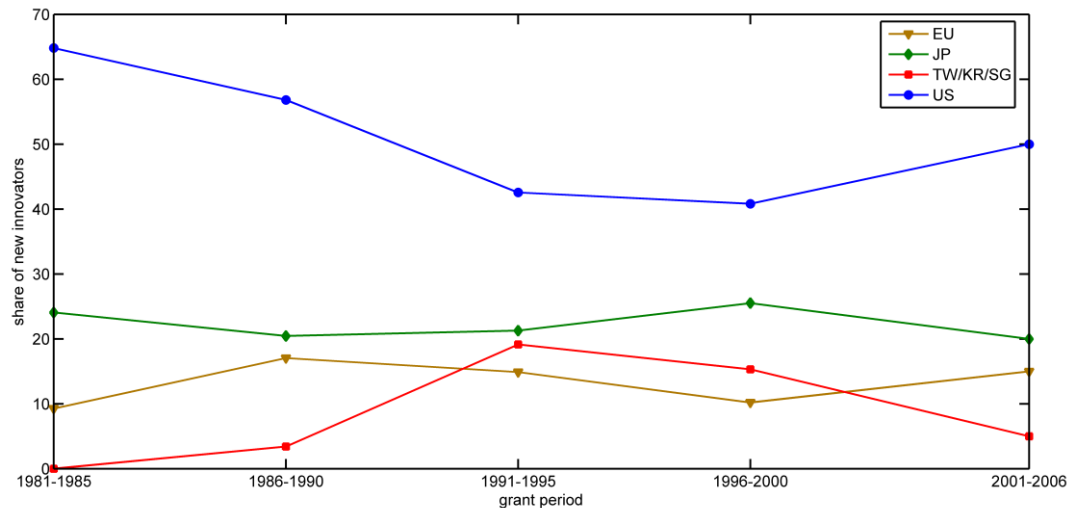


Figure 19: New entrant innovators by country of origin

### 4.3 Theoretical framework

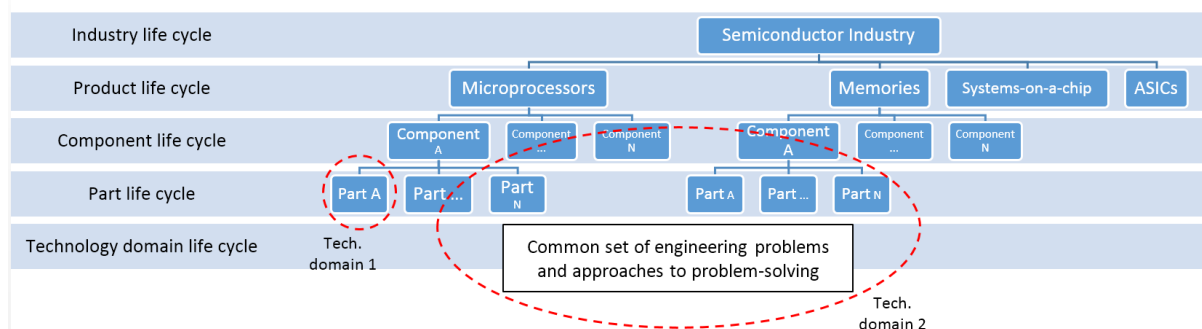
The answers to the questions “*how technology change?*” and “*how are products and industries affected?*” have been seek in a variety of ways. As we have discussed in the introduction of the paper, the prevailing belief is that the emergence of dominant product designs is the selection mechanism that stop the search process and, consequently, reduce the number of players in the competition arena. This allows focusing innovative efforts to process innovation, which, by making the product cheaper, spark diffusion. Eventually the fossilization of product design constrains the generation of novelty and lead to the emergence of decreasing returns to adoption. This eventually increases the probability that the dominant design is re-thought or abounded in search of the new product to introduce.

This is an accurate representation of the life cycle of a single product and there is evidence that the pattern of *market* entry and exit in the industry is consistent with the predictions of the product life cycle theory (Klepper, 1996; 1997). Yet its application to industries with a large heterogeneity of products, some of which are highly customized and based on highly modular technologies, is limited by conceptual difficulties. Semiconductor devices are made by several independent components. Some of them contribute to different products and therefore the underlying technical problems and the way solutions are seek affect the life cycle of several products. Second, in high-tech industries long-run market survival depends on technical capabilities (Lee and Lim, 2001). Therefore *innovative* entry (and exit), defined as the ability to tap in the right technological trajectory, is more informative of a firm’s long-run success than market entry, which could be due to a transitory cost-advantage, in particular for catching-up firms. In an industry characterized by a multi-technology space, the ability of persistently come-up with inventions that proves to be on the right direction with respect to the future technological developments depends on direction of technological change, depends on the life cycle of technology domains. We therefore claim that, for industries like Semiconductors, life cycles should be studied at the technology level.

We define a technology domain as an area of engineering research bounded by a set of common design problems and by similar approaches to problem solving. We adapt Murmann and Frenken’s (2006) nested hierarchy approach to theoretically link the technology domain level to the product level. This theoretical exercise does not have to be seen as an attempt to

formally build a correspondence table between products and technologies. The goal of this work is to illustrate the systemic nature of technology and the relationship between industry, product and technology domains' life cycle. This is used to stress how the most insightful unit of analysis is the technology level. Our theoretical framework is illustrated in Figure 20.

Source: Author's adaptation of Figure 3 from [Murmman and Frenken \(2006\)](#)



**Figure 20: Nested hierarchy of life cycles**

An industry is made by a collection of products. There is general agreement in the recent economic (Murmman and Frenken, 2006), engineering (de Weck et al., 2011) and complex system (Arthur, 2009) literature in describing products as systems of nested hierarchies made by layers of components and parts. The degree of modularity of the system determines whether components can be designed and produced in isolation from each other's. Technology domains can span them both vertically and horizontally or they can be confined to a given component or product. This ultimately depends on the generality of the underlying engineering problem. For instance, miniaturization or reduction of energy-consumption, are very general and ubiquitous problems. The former carries a variety of related sub-problems like velocity saturation or degradation due to overheat, as the technology scaling reaches channel lengths less than a micron. These problems are not isolated and related to a single product or component. In contrast, they affect the whole system. Change at the domain level propagates in the system along multiple paths, generating positive feedbacks or creating cascades of design problems, as shown by Giffin et al. (2008). Therefore, the search for solutions to key design challenges ultimately affects the life cycle of components and products. Consequently, incumbents' and new entrants' innovation prospects depend on their technical capabilities and their knowledge upgrading paths measured at the domain level. Mapping the hierarchical structure of the system goes beyond the scope of this work. For our purposes, we only need to perform the following tasks:

- 1- Identify technology domains;
- 2- Assess their life cycle stage;
- 3- Investigate whether incumbents' and new innovators' revealed technological advantage significantly differs along the life cycle of technology domains.

The technical aspects of the methodology used to address each of the three tasks are explained in details in the next section. Here we highlight the theoretical motivation behind them. A popular approach concerning the problem of identifying technology domains rely on raw information contained in patent database and on patent classification codes such as the US patent codes or the International Patent Classification (IPC). We believe that the well-known skewed nature of the distribution of patent economic and technical value (Silverberg and Verspagen, 2007; Trajtenberg, 1990) makes it mandatory to clean the data by first identifying influent patents. This is done by applying the Main Path Approach (MPA) that identifies patents that lay on the main

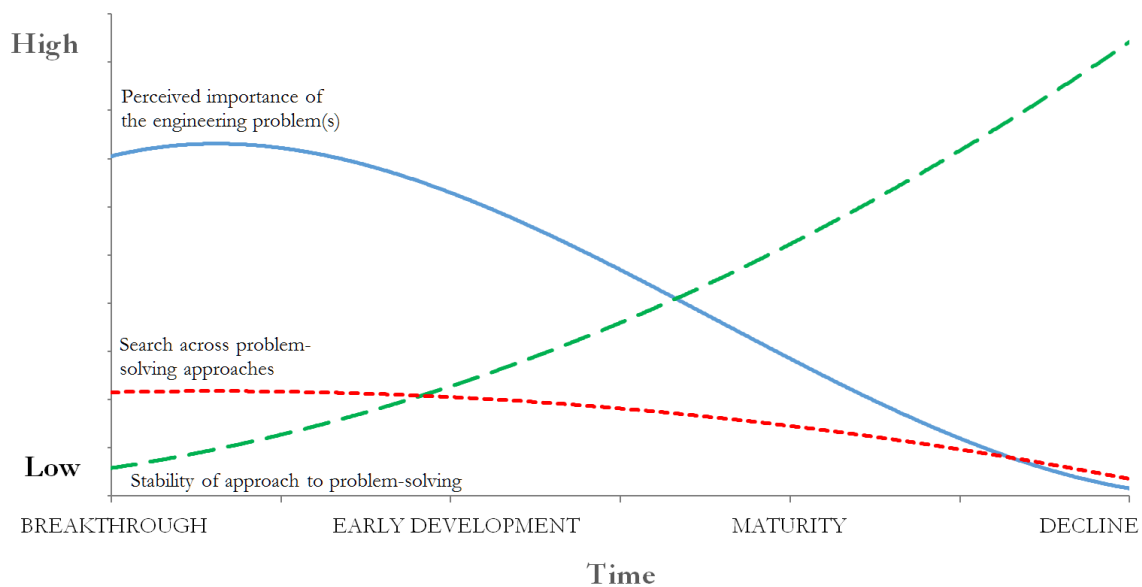
paths of technical improvements, as measured by citation flows. Moreover, we claim that using patent classes to identify domains is not a suitable option given the purpose of this study. Classifications are a very useful source of information but by definition are done ex-post and are subject to a certain degree of subjectivity. For instance, an emerging domain would not be immediately recognized as such. It will require a new classification, which will imply a considerable lag in our ability to recognize emerging technologies. For this reason, we rely on a bottom-up network-based classification that makes use of network community-detection techniques. Domains are identified by clustering patents based on their similarity. The latter is measured by direct and indirect citations. Citations are a proof that the prior-art disclosed in the cited patent has been improved by the cited one. As such, they reveal the similarity of the underlying technical problem and the use of the same problem-solving approach. This perfectly fits with our theoretical definition of domains. Selecting only the most technologically influent patents besides cleaning the data also allows distinguishing innovators from inventors. By the former, we mean players that are able to generate novelty that is later recognized as useful from the point of view of technical progress. The latter are players whose inventive output does not attract sufficient attention to determine the course of technology evolution. In this sense, we use the term innovation in a Schumpeterian way, implying that inventions became innovations only when they are recognized as useful and, therefore, start diffusing. We further distinguish between *incumbent* and *new innovators*. Note that the use of the terms “*new innovators*” or “*incumbent innovators*” rather than new entrants or simply incumbents is purposely made. Industrial organization theory would distinguish between firms that have started *producing* for the first time (new entrants) or have been doing it for a while (incumbents). Since we look at the technological dimension rather than the manufacturing one, we characterize firms by their ability to *generate technological inventions* that lately attracted a significant sequence of engineering improvements.

Once we have theoretically defined what technology domains are, and why we identify them using a bottom-up citation-based approach, we can formalize the theoretical concept of domain life cycle. We argue that the evolution of a technology domain can be described by two variables: the *importance* of the underlying technical problem and the persistence of the *variety* of approaches to problem solving.

An archetypal description of the evolution of a technology domain is presented in Figure 21. Let us suppose that the origin of a given technology domain is a *breakthrough* innovation. These innovations bring a completely new set of engineering problems that are very loosely related with previous solutions. The problem-solving approach is therefore disconnected with past experience. This implies that a variety of search strategies is applied to seek the solution. Breakthroughs are obviously rare and are usually identified as such only ex-post. Our approach identifies potential breakthroughs ex-ante as clusters of related problems with no or loose connections with the past that attract a lot of innovative effort by some of the players. In other words, finding solutions to these problems is considered as an important task. If the underlying problems are recognized as important and a solution is possible, at some point the variety of search strategies starts to decrease. Problem solving begins to be path-dependent and the persistence of a common approach increases. The underlying problems are still considered as important but they attract slightly less innovative efforts than before. The domain moved toward its *early development* phase. As time goes by the domain enters its *maturity* phase, a bulk of existing knowledge accumulates, the search of alternative approaches greatly reduces and the problems themselves become to be perceived as less important. This can be due to at least three reasons. First valid solutions have already been found (i.e. technological progress have moved further). Second, technological development at the product level has taken a different trajectory and oth-

er engineering problems are perceived as bottlenecks of progress. Third, the problems start to be perceived as unsolvable. In any of the three cases if a general agreement on the reduced importance of the problem emerges the domains move to its decline or *exhausting* stage. Innovative efforts drop dramatically and the remaining gleams of inventive activity, if at all, follow clearly predefined problem-solving approaches. This destiny is not ineluctable. Some players might think that searching for better solutions is still worthy, perhaps because of a different vision of the future development of the technological trajectory or because of the attempt to improve older generations of a given product or technology. This is likely to be the case for players engaged in technological catching-up endeavours. When this happens, there is a renewed interest in the set of technical problems and a revamp in the search of alternative approaches. The domain enters into a *renewing* phase. This type of life cycle is portrayed in Figure 46 in the Appendix A.4.1. If the renewal phase is successful and the new approaches are promising a new life cycle might start, otherwise the domain might face decline anyhow.

When we described the archetypal life cycle of a given technology domain we assumed it started with a breakthrough. Besides a successful renewal of an old domain, another exception to the breakthrough kick off exists. A life cycle might be initiated by the emergence of *disruptive* technology domains. Christensen (1997) defined disruptive technologies as those that initially perform worse than the current best practices and address a different market but eventually outperform current technologies even in their own market. We use the word disruptive to describe domains whose engineering problems initially do not attract much innovative effort because they are not generally recognized as important. Furthermore, in these domains, the search of solutions follows unconventional paths. However, if promising approaches to problem solving arise, the importance of the problem and the value of the new design approaches might be recognized by many players. These domains would then start attracting more inventive effort. Eventually, this would spark the life cycle. This is shown in Figure 47 in the Appendix A.4.1.



**Figure 21: Archetypal life-cycle of a given technology domain starting with a breakthrough**

The exploration of different approaches to a problem, has a clear theoretical relationship with the concept of technological trajectories. A technological trajectory is defined by Dosi (1982) as the direction of problem-solving activities within a technological paradigm. Yet, although conceptually related there is an important difference concerning the level of analysis and the way

they affect technological progress. Trajectories are typically defined at the product level. They are the results of design choices on which features of the product to improve, in particular when these features are affected by trade-offs (e.g. computational power vs. energy consumption). Technological trajectory affect and are affected by the life cycle of technology domains. On the one hand, choices along the trajectory obviously imply that some engineering problems will be perceived as more important than others and consequently attract more innovative efforts. Depending on the novelty of the problem, the urge to find appropriate solutions will either spark a variety of search strategies or follow predefined and more conservative approaches. On the other hand, the solution of problems that affect several components and/or products, pushes innovative efforts toward some products features rather than others or might even allow braking the trade-off.

There is also a clear relationship between the life cycle of technology domains and catching-up strategies followed by latecomers. Lee and Lim (2001) defined three types of catching-up: *path-following*, *stage-skipping* and *path-creating*. When the latecomer firm just follows the same path taken by the forerunner (with a narrowing delay), the catching-up process is said to be path-following. When, instead, the latecomer firm learn so quickly to be able to skip one or more generations of the technology, catching-up follows a stage-skipping pattern. Finally, the authors define path-creating catch-up. This is defined as the situation in which the process of learning and assimilation of older generations of a given technology, by a latecomer firm, results into significant technical improvements that take a different direction compared to the current path followed by leaders. The authors argue that stage-skipping and path-creation are better described as leapfrogging rather than catching-up as they involve doing something different from what previously done by the leaders. There is a strong analogy between the life cycle stage of a given technology domain and the type of catching-up followed by latecomers. Successful path-following catching-up would correspond to initially specializing in exhausted areas and then systematically move backward along the life cycle, specializing in mature, early-growth and emerging areas at each subsequent time. If any of the steps would be skipped along the catching-up process than we could describe it as a stage-skipping type, or leapfrogging. Taking Lee and Lim's definition literally, path-creating would correspond to an early specialization in breakthrough or disruptive areas, as it reveals that the latecomer is exploring its own path. However, we claim that specialization in renewing areas also falls into the path-creating category of catching-up given the explorative nature of the learning endeavour.

#### **4.4 Methodology and Preliminary Data Analysis**

To empirically test the validity of our theoretical framework, identify semiconductor technology domains' and analyse players' revealed technological advantage over the life cycle, we need to perform a number of practical tasks. Table 7 reports these tasks together with the methodology used to address them. The table is intended to serve as a roadmap to navigate the next subsections, where we systematically explain the methodology used to tackle each task.

**Table 7: Methodological roadmap to analyse the life cycle of technology domains and specialization patterns of new and incumbent innovators**

| Task  | Methodology used                                       |
|---|--|
| Define a method to identify technologically influent patents  | Network of Main Paths Approach                         |
| Define a method to cluster patents around relatedness of problems and approaches to problem-solving (i.e. identify technology domains)    | Community detection                                    |
| Measure importance of the underlying problem and persistence of the problem-solving approach  | Communities' patent composition analysis               |
| Define boundaries of the importance and persistence to categorize life cycle stages of technology domains, from breakthrough to declining | Analysis of the distribution of patent types           |
| Analyse which actors (i.e. new or incumbent innovators) are more active in which stages.  | Analysis of the Revealed Technological Advantage Index |

#### 4.4.1 Identification of technologically influent patents

Patent technical and economic value is highly skewed (Silverberg and Verspagen, 2007; Trajtenberg, 1990). Only a few patents stand out of the crowd in terms of their importance for the course of technical change and their economic value. Therefore, there are two reasons to reduce population of patents to a sample of the important ones. First marginal patents and their citations create noise that can make it difficult to identify important signals from the data. Second, the reduced network size is computationally convenient for subsequent data analysis. The *Network of Main Paths* (NMPs) is a methodology developed to identify the routes through which knowledge diffuses in large citation networks made of patents or publications (Martinelli, 2009; Fontana et al., 2009; Verspagen, 2007). When applied to patent citation networks this methodology allows analysing the evolution of the main sequences of technological improvements in a given industry or technological area. The first building block of this approach relates to the meaning of patent citations. If patent B cites patent A then the former improves upon the latter. In other words, A represents the state-of-the-art concerning the particular technology described in patent B at the moment in which patent application B was filed. Therefore citations can be interpreted as a measure of technological relatedness<sup>14</sup> and provide insights on the direction of technological change. In particular, the citation relationship proves that the citing patent focused on the same engineering problem than the cited one. Furthermore, it also highlights the use of a similar approach to problem solving as otherwise the relationship between the technical claims of citing and cited patents would not be justified. In this case, the latter would not be considered as the relevant prior-art for the former. Obviously, a patent can cite and be cited by many other patents. Hence, if we want to follow the main trajectories of technology evolution among a set of

<sup>14</sup> From this perspective, the well-known fact that many, if not most, of the citation are added by the patent examiner rather than the applicant plays in our favor. Indeed patent applications are examined by expert in the field of the technology described by the patent. Therefore, citations added by examiners can be seen as an even more objective measure of technological relatedness among patents. Obviously examiner-added citations are instead much more of a problem if one wants to use them as a measure of spillover between patent assignees. This is not the case for this work.

patents, we first need to decide which direction to take at every junction. This is what the algorithm to define the NMPs does. First, we calculate the weight of every citation using the *search path node pair* (SPNP) algorithm, developed by Batagelj (2003) based on the original measure introduced by Hummon and Doreian (1989). The SPNP returns the number of times that each citation link lies on all possible paths connecting any node to anyone else<sup>15</sup>. This is easily calculated by multiplying the number of patents that reach (through direct and indirect citations) the cited patent by the number of patents that are reached (directly or indirectly) by the citing patent. Therefore, a high SPNP weight indicates that the given citation and the two patents involved are located in a highly connected and connecting area of the network. This means that the given citation has a strong technological influence, as many paths of technological improvement pass through it. The NMPs is then identified by following the paths emanating from start nodes (nodes that are cited but not cited), taking at each junction the direction of the citation which carries the highest weight, till an end point (a node who cites but is not cited) is reached. This process can be repeated several time by cumulating windows of time, (e.g. from time  $t$  till  $t+1$ , then from  $t$  till  $t+2$ , and so on). By computing the NMPs for each period we can observe how the entrance of young patents at each point in time affects the presence of old ones in the network of main paths (i.e. the persistence of old technological trajectories). When newly granted patents connect to previously well-connected patents, technical improvements follow the same paths of citations of the previous period(s). In this case, the technological trajectories are said to be stable and cumulative. We interpret this case as an instance of stability of problem-solving approaches. To the contrary, if the new patents connect to paths that were previously underexploited the patent composition of the NMPs changes and the technological trajectories are affected by a discontinuity. We interpret the latter as a case of search of alternative problem-solving approaches. In a nutshell a given NMPs at each point in time is populated by three types of patents: those recently granted (i.e. *young patents*), older patents that appeared in previous snapshots of the NMPs (i.e. *old persistent patents*) and older ones that show-up for the first time (i.e. *old new patents*). The distribution of patents across the three types is instrumental in identifying the life cycle of technology domains, as we will explain in Section 4.3.

We apply the NMPs methodology to the whole citation network of semiconductor technology-related patents granted by the USPTO between 1976 and 2006. First, we extract all US patents belonging to the following five US technology sub-classes: 438-process, 257-product, 326-materials, 505-programmability, 716-design. Then we create the citation network and extract the largest connected component. The latter is used to feed the NMPs algorithm that extract the most important paths of citations based on the SPNC weights and identify the patents laying on them. The largest component of the resulted reduced network is composed by the set of patents that we claim being the most influential from the perspective of technical progress. Table 8 reports the network size at each layer of data reduction. Figures showing the main component of the NMPs for the six periods are reported in the Appendix A.4.2. The technology domains are highlighted in different colours (these areas have been identified through the community detection procedure explained in Section 4.2).

---

<sup>15</sup> Readers familiar with recent developments in graph theory will recognize the similarity with the measure of random-walk betweenness centrality (RWBC) introduced by Newmann (2005). Indeed SPNP and RWBC are essentially the same measure. The former had been defined by Hummon and Doreian (1989).



Table 8 shows that the main component of the NMPs for the periods 1976-1995 and 1976-2006 decreased in size compared to the periods before them. Both were analysed in the previous chapter. The drop in the period 1976-1995 has been explained by the action of new innovators from latecomer countries that happened in the mid-1990s. Their learning strategy consisted in searching for alternative-approaches to the existing sets of engineering problems. However the disruption that they caused to the main paths was temporary. Some patents from the main component of the NMPs in 1976-1995 moved to the second component in the next period. Yet they were brought back into the largest component in 1976-2000, revealing successful attempts to make a synthesis of different problem-solving approaches. In the context of this paper, it is more interesting to focus on the drop in the size of the main component of the NMPs that occurred in the last period (1976-2006). This second drop was also discussed in the previous chapter. In this case, the reason was a change in the ranking of engineering problems. At the beginning of the 2000s, newly granted patents connected more to the second largest component of the NMPs than to the first. This means that a change in the ranking of engineering problems' importance occurred in this period. Priority of innovative effort shifted from those related to domains pertaining to the largest component to those found in the second one. The analysis of patents titles and abstracts revealed that technology domains found in the second component focused on engineering problems related to LCD displays, in particular for e-readers and flat televisions. This suggests that the second largest component of the NMPs is composed of domains more related to entertainment and portable devices than to desktop computers and laptops. What we observe in this period could therefore be a case of overlap between the life cycle of products and technology domains. Given the importance of engineering problems related to the second component of the NMPs in the last period under observation, we include it in the analysis performed in the rest of the paper.

**Table 8: Basic network statistics**

|  | 76-80 | 76-85 | 76-90 | 76-95  | 76-00  | 76-06  |
|--|-------|-------|-------|--------|--------|--------|
| Whole network - number of patents                            | 2079  | 5631  | 12533 | 26853  | 54086  | 114097 |
| Whole network - number of citations                          | 2712  | 13310 | 40255 | 102957 | 272843 | 779076 |
| Main component - number of patents                           | 1703  | 5385  | 12348 | 26686  | 53874  | 113756 |
| Main component - number of citations                         | 2469  | 13164 | 40145 | 102864 | 272728 | 778890 |
| Network of Main Paths - number of patents                    | 1445  | 3490  | 6042  | 10107  | 15387  | 23428  |
| Network of Main Paths - number of citations                  | 1403  | 3291  | 5697  | 9489   | 14588  | 22077  |
| Network of Main Paths - Main Component – number of patents   | 694   | 1540  | 2678  | 2043   | 4557   | 3544   |
| Network of Main Paths - Main Component – number of citations | 756   | 1597  | 2734  | 2064   | 4617   | 3562   |

#### 4.4.2 Grouping technologically influent patents into technology domains

What we have analysed so far is just the size of the NMPs. We need to identify technology domains in this network. Given our definition of technology domains as areas of research characterized by commonality of problems and approaches, the best way of identifying them is by community detection. It became a common practice to analyse large networks' community structure in order to split them into partitions. Partitional and agglomerative hierarchical clustering

methods have been defined to identify such structure. We use a method proposed by Newman (2004) based on the concept of modularity. Modularity is defined as the fraction of links (citations in our case) in the network that falls within a community. The algorithm maximizes modularity. This allows identifying communities as areas of the networks whose nodes are more related to each other's than with nodes outside the community. Technical details about the Newman's community detection algorithm can be found in the Appendix A.4.3, where we also validate the quality of the algorithm's results. We chose to use the Newman algorithm because, contrary to other popular community detection algorithm like, for instance, the Newman and Girvan one (2003), the former provides a benchmark to evaluate the quality of the partition and does not require to arbitrarily choose the number of communities to be identified. Indeed the modularity maximization procedure and the comparison with equivalent random networks returns the best partition of the network analysed, without assuming a pre-existing community structure.

Some basic statistics about semiconductor technology domains identified by Newman's algorithm are reported in

Table 9. The high values of modularity (always higher than 0.85) reveal a strong underlying community structure within the largest component (and the second one in the last period) of the NMPs. This provides empirical support for the existence of several, relatively separated, areas of research within the semiconductor industry. The algorithm identifies a number of domains varying between 14 and 15 over the periods observed. The size of the largest area changes quite a lot. So does the standard deviation and the coefficient of variation.

**Table 9: Basic statistics for the technology domains identified by Newman's algorithm**

|                                      | 76-80  | 76-85  | 76-90  | 76-95  | 76-00  | 76-06<br>(1st Comp.) | 76-06<br>(2nd Comp.) |
|--------------------------------------|--------|--------|--------|--------|--------|----------------------|----------------------|
| Number of patents                    | 694    | 1540   | 2678   | 2043   | 4557   | 3544                 | 2762                 |
| Modularity                           | 0.8567 | 0.8789 | 0.9013 | 0.9066 | 0.9161 | 0.9021               | 0.8967               |
| Number of domains                    | 14     | 15     | 14     | 14     | 15     | 15                   | 14                   |
| Size of the main domain              | 128    | 328    | 368    | 272    | 637    | 701                  | 489                  |
| % of patents in main domain          | 18,44% | 21,30% | 13,74% | 13,31% | 13,98% | 19,78%               | 17,70%               |
| Size of smallest domain              | 15     | 29     | 52     | 65     | 62     | 73                   | 53                   |
| % of patents in smallest domain      | 2,16%  | 1,88%  | 1,94%  | 3,18%  | 1,36%  | 2,06%                | 1,92%                |
| Average cluster size                 | 49,57  | 102,66 | 191,29 | 145,93 | 303,80 | 236,27               | 197,29               |
| St.dev.                              | 34,16  | 80,38  | 80,41  | 69,76  | 143,03 | 149,51               | 118,04               |
| Coefficient of variation (St.dev/Av) | 0,69   | 0,78   | 0,42   | 0,48   | 0,47   | 0,63                 | 0,60                 |

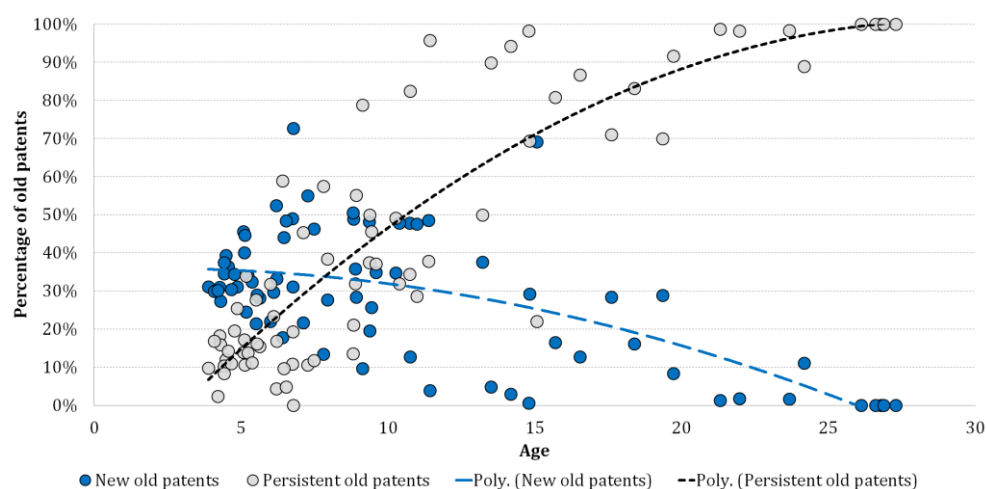
The large size differences among technology domains hint to the importance of analysing their life cycle. In the next subsection, we explain how we identify the life cycle stages of technology domains that we have just identified.

#### 4.4.3 Characterizing technology domains according to their life cycle stage

Our method to identify the life cycle of technology domains is based on the existence of three types of patents that are found in the NMPs at each point in time: *young*, *persistent old* and *new*

*old*. Young patents are those granted in the last period of observation. Persistent old patents are those that have already been part of the largest component of the NMPs at least once in the periods before the one observed. In our analysis, we focus on six periods: 1976-1980, 1976-1985, 1976-1990, 1976-1995, 1976-2000 and 1976-2006. Let us take, for instance, the last period 1976-2006. For this period, the three patent categories can be described as follows. Young patents are those granted after the end of the previous period (i.e. from 2000 till 2006) which connects to the main component of the NMPs. Persistent old patents are those who showed up in the main component of the NMPs at least once in one of the previous five periods. New old patents are those granted before 2001 which had never been part of the main component of the NMPs before. The distinction between persistent old patents and new old patents allow us to distinguish domains where there is no search of alternative approaches, from those who are exploring a new path. Furthermore it also help us to differentiate between areas which are young but nevertheless building on previously explored technological paths and young areas which are not related to any technological solution that have been developed in the past. Figure 22 shows the relationship between the type of old patents and the age of the technological areas. Each circle stands for one of the technology domains identified over the six periods. Its position on the horizontal axis reflects the age of the area. The vertical axis coordinate is given by the percentage of old new and old persistent patents found in the domain (each domain counts for two circles in Figure 22). Dashed lines are lines of best fit obtained by linear regression using a second degree polynomial as mathematical model.

The figure shows that young domains are more likely to build on previously unexploited technological solutions (new old patents) than known ones (persistent old patents). Therefore, search across possible problem-solving approaches is higher. To the contrary, the more a domain grows old, the more likely it will follow a stable and previously defined approach to problem solving. The two curves closely resemble the patterns sketched in Figure 22. This confirms our theoretical predictions based on the cumulative nature of technological change. Figure 22 also clearly shows that patent composition within a technological area changes drastically with age. Our classification method follows the intuition that it is possible to categorize domains' life cycle stages based on the relative number of young, persistent old and new old patents, they are composed of. This allows defining all the stages of the life cycle of technological areas, from emerging to declining.



**Figure 22: The relationship between persistent old patent, new old patents and the age of semiconductor technology domains**

Based on the theoretical framework discussed in Section 4.3 and the empirical findings shown in Figure 22, we propose the following theoretical correspondence between each life cycle stage and the patent composition that reflects it.

### *Breakthrough*

Breakthroughs break the usual pattern of knowledge cumulativeness that normally characterizes technical change. Their relationship with previous solutions is very little if existent. We argue that domains in their breakthrough stage are characterized by a large number of young patents and a few new old and persistent old patents if at all.

### *Disruptive emerging areas*

We argue that disruptive technological areas are characterized by the presence of several young patents that builds largely on previously disconnected patents and very little on persistent old ones. This reflect the high search across possible approaches to problem solving which characterize emerging areas but also the peculiar focus on previously unexploited existing solutions which make the domain disruptive in nature. The other marking trait of disruptive domains is that the underlying set of problems initially does not receive much attention. The latter two characteristics distinguish disruptive domains from breakthroughs.

### *Early development*

If successful, disruptive or breakthrough domains are developed further and move to a stage of early growth. During this stage, the attractiveness of the area of research is high and the technological trajectory starts to consolidate. Therefore, the number of young patents is high, the presence of persistent old patents increases and the one of new old patents decreases.

### *Maturity*

Maturity is similar to the early development stage with the only difference that the domain now attracts much less innovative efforts (i.e. fewer young patents connect to it) and technological change becomes increasingly cumulative. This means that the number of persistent old patents keeps growing, to the detriment of the exploration of alternative approaches.

### *Renewing*

After the maturity stage the evolution of a given technology domain is at a crossroad. The development of the given technology could be either stopped or get new vigour. In the former case, the domain begins exhausting. In the latter, it enters into a renewing stage. In this case, alternative paths are explored to avoid obsolescence. This might begin a new life cycle or just extend the life of a technology domain for a short while without avoiding its imminent decline. The renewing stage is characterized by a few young patents that build extensively on new old ones and on some persistent old patents.

### *Exhausting*

Exhausting (or declining) areas are characterized by very few, if any, young patents, a large number of persistent old patents and almost no new old ones.

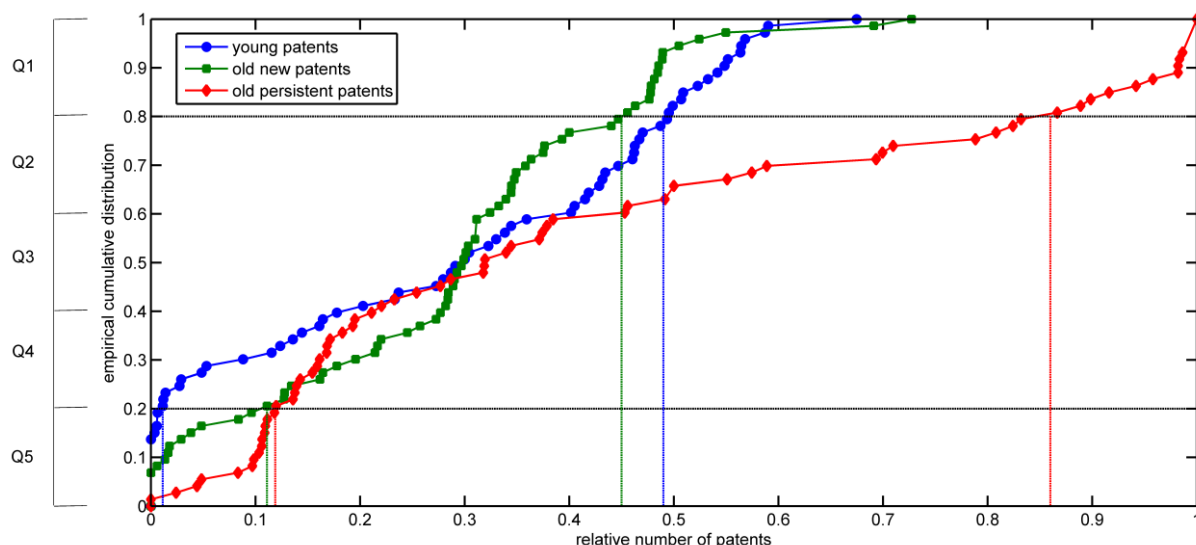
At this point, we have a theoretical definition of the life-cycle stages of technology domains and the preliminary characterization of them according to the relative number of young, old persistent and old new patents that is found in each domain. To make our methodology operational we need a practical way to formally distinguishing one stage from the other. Consider a triangu-

lar shaped space in which the horizontal axis measures the relative number of old persistent patents in a given domain and the vertical axis reports the relative number of young patents. The structure of the space is such that domains can only locate in the lower triangle that is defined by the axis and the diagonal connecting the maximum values of the two axis (i.e. 100). This is because the relative number of patents per each category is constrained between 0 and 100. Therefore, by construction the orthogonal distance of each domain from the diagonal measure the relative number of old new patents. We call such space the *life-cycle space of technology domains* as the entire life of a given domain can be described by movements along this space. The space is reproduced in Figure 24. However, before to discuss the figure, let us first explain step-by-step the process behind its creation. We first need to draw borders on such space that will help us identifying the areas corresponding to each life-cycle stage. To accomplish this task we need to formally quantify the relative number of young, persistent old and new old patents that a domain must have for its life cycle to be in a given stage. Quantify how much is a lot is a task that is best done by comparison. Therefore, we first take all domains identified by Newman's algorithm over the periods 1976-1985<sup>16</sup>, 1976-1990, 1976-1995, 1976-2000 and 1976-2006, we look at the percentage of young, persistent old patents and new old ones in each area and then we plot the distribution of these percentages. This is shown in Figure 23, where each of the domains is split into three observations indicating the percentage of young, new old and persistent old patents it is composed of. On the horizontal axis, we have the values for the percentages of each category of patents that are part of one of the technology domains, whereas on the vertical axis we have the cumulative percentage of the distribution, meaning the percentage of observations with a value smaller than the value on the horizontal axis. We drew two horizontal dashed lines to clearly separate the top 20 percent from the mid-60 percent and the bottom 20 percent of the distribution. This allows us to identify the border values for the first quintile and the last quintile. For instance, if we look at the distribution of the relative number of young patents among all technology domains we see that 20 per cent of the domains have less than 1.14 per cent of young patents, 60 per cent have between 1.14 per cent and 49.35 per cent of them and 20 per cent have more than 49.35 per cent of young patents. For instance, this means that a given domain can be said to have many young patents if more than 49.35 per cent of its patents are young. In this case, the remaining 50.65 per cent is distributed between new old patents and persistent old ones. The same exercise can be applied to new old patents and persistent old ones. In the former case 20 per cent of the domains have less than 11.11 per cent of new old patents, 60 per cent have between 11.11 per cent and 45.57 per cent of them and 20 per cent have more than 45.57 per cent of young patents. Finally, if we look at the distribution of the relative number of persistent old patents we see that 20 per cent of the domains have less than 11.97 per cent of them, 60 per cent have between 11.97 per cent and 86.67 per cent and 20 per cent have more than 86.67 per cent. It is important to notice that there are no domains purely composed by young or new old patents. Nevertheless, a few are entirely made of persistent old patents. From a NMPs methodological point of view we can argue that a domain purely made by young patents or by new old ones would be disconnected from the main component of the NMPs by construction and therefore not observed. To the contrary, albeit rarely, domains entirely com-

---

<sup>16</sup> We cannot use the first period, 1976-1980 because, being the initial period, by construction all the areas are entirely composed by young patents.

posed by persistent old patents can be found in the main component of the NMPs. They indicate technological ancestors upon which newer solutions build.



**Figure 23: Empirical cumulative distribution of the percentage of young, new old and persistent old patents for all the domains in the periods 1976-1985, 1976-1990, 1976-1995, 1976-2000 and 1976-2006**

Now that we have more precise quantities of young, new old patents and persistent old ones, we can use them to elaborate a more precise definition of the life cycle stages of technological domains. Table 10 reports the thresholds that define the amount of each type of patents to be found in a given domain for it to be classified in one of the life cycle stage reported in the left column. We call this thresholds *quantile borders*. For instance, for a domain to be classified as a breakthrough it needs to have at least 49.35 per cent of young patents, less than 45.57 per cent of new old ones and less than 11.97 per cent of persistent old patents.

However, the quantile borders alone are not sufficient to determine life cycle stages. The main reason is that, being thresholds, quantile borders suffer from the drawback that areas that lay very close to the border might actually be more similar to the areas located on the other side of the border than to the other areas located on the same side. This problem is similar to the one of defining homogeneous groups of people living in areas whose borders have been set on paper, without considering the common characteristics of people living close to the border. In other words, we would like to have borders that respect the geography of the life-cycle space. Therefore, the initial quantile borders are used to calculate centroids, which serve as *basins of attraction*. To sum up, first we calculate the quantile borders for the distribution of the percentage of young, new old and persistent old patents for all the domains in the periods 1976-1985, 1976-1990, 1976-1995, 1976-2000 and 1976-2006 (Table 10). Then we use them to preliminary identify regions of the life-cycle space that are coherent with the theoretical description of the life cycle stages of technology domains and the empirical distributions of young, persistent old and new old patents. Afterwards we calculate the centroid for each of the preliminary defined areas of the life-cycle space. Finally, we compute the distance to each of the centroids for each technology domain identified through Newman's algorithm. The life cycle stage of each technology domain is then identified by assigning each domain to the closest centroid. This procedure is shown in Figure 24.

**Table 10: Patent distribution quantile borders by patent type and life cycle stage**

Quantile classification

Many

Q1 (i.e. top 20%)

Mid

Q2, Q3, Q4 (i.e. mid 60%)

Few

Q5 (i.e. bottom 20%)

Quantile borders for the technological area life cycle stages

Young patents

New old patents

Persistent old patents

|                             |                           |                           |                                    |
|-----------------------------|---------------------------|---------------------------|------------------------------------|
| Breakthrough emerging areas | Many = Q1 (>49.35%)       | Few-mid = Q2-Q5 (<45.57%) | Few = Q5 (<11.97%)                 |
| Disruptive emerging areas   | Few-mid = Q2-Q4 (<49.35%) | Many = Q1 (>45.57%)       | Few = Q5 (<11.97%)                 |
| Early growth areas          | Many = Q1 (>49.35%)       | Few-mid = Q2-Q5 (<45.57%) | Mid Q2-Q4 = (11.97% ≤ ... <86.67%) |
| Mature areas                | Few-mid = Q2-Q4 (<49.35%) | Few-mid = Q2-Q5 (<45.57%) | Mid Q2-Q4 = (11.97% ≤ ... <86.67%) |
| Renewing areas              | Few-mid = Q2-Q4 (<49.35%) | Many = Q1 (>45.57%)       | Mid Q2-Q4 = (11.97% ≤ ... <86.67%) |
| Exhausting areas            | Few = Q5 (<1.14%)         | Few = Q5 (<11.11%)        | Many = Q1 (>86.67%)                |

Each node stands for one of the technology domains identified in Section 4.2. The size of the node is proportional to the size of the given domain measured by the number of patents. The location of a domain on the life-cycle space is informative of its patent-composition and therefore of its life-cycle stage. In Figure 24 red lines highlight quantile borders reported in Table 10 and centroids are marked with a red 'x'. Domains that share the same colour fall within the basin of attraction of the same centroid. This means that they are closer to that centroid than to any other one and therefore are in the life-cycle stage indicated by the centroid. Note that by connecting centroids of subsequent life cycle stages and tracking the evolution of the relative number of young, old persistent and old new patents, curves similar to those reported in Figure 21 emerge. This highlights the strong connection between the theoretical description of the life-cycle of technology domains and the methodology used to trace it.

Now we have a classification of the life cycle stage of each technology domain. To test its logical consistency we trace movements from each life cycle stage to the other ones. Of course, for our classification to be coherent, we should observe movements consistent with time. This means that, for instance, patents that are classified into a technology domain in its early development stage in period T should be mainly part of a domain classified as mature in the next period. Some might still be found in an early-development stage. This would indicate that the life cycle of that domain is relatively slow. Some others might jump over stages and be found in renewing or exhausting domains. This would indicate that the life cycle of that domain moved faster in the period observed. The crucial aspect is that they should not be found in large numbers in an earlier stage, otherwise the time consistency of our methodology would be broken. A small number of patents could actually move back to an earlier stage but this can only happen when some patents from one domain serve as foundation for a younger one in the next period. This possibility is intrinsic to the evolution of communities as defined by Newman's algorithm and the network of main path approach. However, this cannot happen in large numbers because oth-

erwise the new domain would not be younger than the original one and would then be classified in the same life cycle stage than the latter, or in one of the followings.

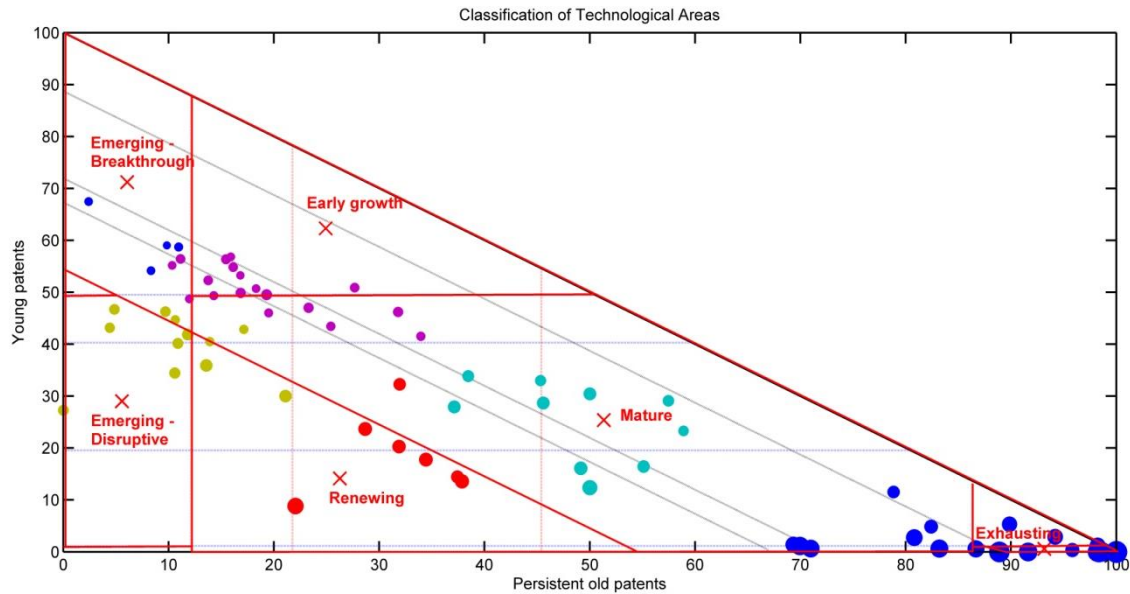


Figure 24: The life-cycle space

Table 11 shows how many patents from domains which, in period T, were in one of the life cycle stages listed on the rows moved, in the next period, to any of the domains whose life cycle stage in T+1 is indicated in the columns.

Table 11: Movements from one life cycle stage to the others over consecutive periods

|              | Breakthrough | Disruptive | Early growth | Mature | Renewing | Exhausting |
|--------------|--------------|------------|--------------|--------|----------|------------|
| Breakthrough | 0.00%        | 24.36%     | 50.00%       | 6.41%  | 16.67%   | 2.56%      |
| Disruptive   | 1.99%        | 4.42%      | 25.88%       | 24.56% | 29.65%   | 13.50%     |
| Early growth | 0.39%        | 1.49%      | 14.48%       | 29.89% | 15.88%   | 37.87%     |
| Mature       | 2.00%        | 1.80%      | 4.60%        | 12.00% | 6.00%    | 73.60%     |
| Renewing     | 0.00%        | 0.77%      | 0.77%        | 3.84%  | 9.72%    | 84.91%     |
| Exhausting   | 0.00%        | 0.00%      | 0.13%        | 0.00%  | 0.00%    | 99.87%     |

The table clearly proves that our methodology is logically consistent as most of the patents follow the expected movement to “older” life cycle stages (to the right of the diagonal) and very few moves to “younger” domains whose life cycle stage is antecedent the one of origin (to the left of the diagonal). Having proved the consistency of our methodology, we can now introduce the answers to the paper’s research questions.

#### 4.4.4 Measuring comparative technological advantage along the life-cycle

In the introduction of our paper, we raised two research questions about the role played by incumbent and new innovators along the life-cycle of technology domains. In order to analyse aggregate comparative technological advantage we propose an original index that returns a macro-aggregation of micro-comparative technological advantage of individual firms. Our specialization index, which we call *SPEC*, builds on the well-known revealed technological advantage index



(RTA). The RTA is a specialization index defined by Soete (1987), which builds on the Ricardian concept of comparative advantage and, more precisely, on the revealed comparative advantage index as defined by Balassa (1965). The intuition behind the RTA is that even if a given entity (countries, firms, geographical regions) might have less patents than other entities in absolute terms, there might still be areas of technology in which it enjoys a comparative advantage. This means that such entity could be able to produce comparatively more patents in a given technological area than in the overall industry. Indeed, the index reveals the domains in which a given entity performs comparatively better. This reflects the entity's comparative advantage in terms of research productivity in those domains. Neoclassical economic theory would suggest that agents (firms or countries) should specialize in those domains where they enjoy comparative advantage. Obviously, this is a static suggestion that does not take into account the possibility of knowledge upgrading. Our use of the SPEC index is intended to investigate in which life-cycle stages agents' capabilities significantly differ, in particular between new and incumbent innovators. However, it must not be understood as a suggestion that agents should necessarily specialize in those domains permanently. To the contrary, in Section 4.5.2, we seek evidence of knowledge upgrading by looking at how the revealed comparative advantage changed over time. The original version of the RTA index is calculated as follows:

$$RTA_{ik} = \frac{x_{ik} / \sum_i x_{ik}}{\sum_k x_{ik} / \sum_{i,k} x_{ik}} \quad (4.1)$$

Where  $x_{ik}$  is entity's (country or firm)  $i$  number of patents in domain  $k$ . The RTA index is equal to zero when entity  $i$  holds no patents in the given domain  $k$ . When the index is equal to 1 entity  $i$ 's patent share in area  $k$  is equal to its share in all areas. Values of the index greater than 1 indicate comparative advantage in the given domain. The original version of the index is not symmetric, meaning that it is bounded to zero for comparative disadvantage in the domain but unbounded for comparative advantage. This causes problems when the RTA is used in econometric models or when one wants to compare the shape of its distribution for different entities. Since in this work we intend to do the latter we therefore opt for the symmetric version of the RTA (SRTA), which is calculated as follows:

$$SRTA_{ik} = \frac{RTA_{ik} - 1}{RTA_{ik} + 1} \quad (4.2)$$

In its symmetric version the index ranges from -1 (full negative specialization) to  $\lim_{RTA \rightarrow \infty} 1$  (full positive specialization), with values greater than 0 indicating comparative advantage in the domain.

We use the symmetric RTA as a basis to construct an index that gives a micro-founded picture of specialization patterns at the aggregate level. We first need to estimate the probability density function (pdf) of the SRTA for each country. The pdf returns the probability to observe a given SRTA value if we choose a firm at random out of the sample of firms belonging to a given country. We use a kernel smoothing function to estimate the probability distribution that best fits the empirical (cumulative) distribution of the SRTA for the given entity. The kernel density function estimates the probability to observe a given SRTA for the whole range of the SRTA index (from -1 to 1). This improves our ability to compare entities of different size as the empirical distribution for small entities relies on fewer observations than for large entities. Once we estimated the probability density function, we compute the SPEC index as follows:

$$SPEC_{ik} = \sum_{j=0:0.1:1} SRTA_j \cdot \rho(SRTA_j)_{ik} \quad (4.3)$$

Our specialization index  $SPEC_{ik}$  is the weighted sum of the probability  $\rho$  to observe SRTA values at the firm level reflecting comparative advantage in the given domain (i.e.  $SRTA > 0$ ). Indeed  $\rho(SRTA_j)_{ik}$  is the probability to observe a given SRTA value  $j$  greater than zero (i.e. positive specialization) among the whole sample of SRTA values calculated for the area  $k$  for all firms belonging to the given country  $i$ . This probability is multiplied by the strength of specialization, namely by the value of the SRTA  $j$ , which, ranging from 0 to 1, effectively serves as a weight for the sum. We limit the SRTA range to positive values because we are not interested in comparative disadvantage. In other words, a large value of the SPEC index means that, if we extract a firm at random out of the sample of firms from the given country, that firm has a high probability to be strongly specialized in the area under consideration. It is important to note that our index focus on the right tail of the distribution of SRTA. This is an improvement over traditional approaches that calculates the SRTA at the firm level and then averaged it at the country level. This approach fails to realize that comparative advantages are rarer than comparative disadvantages. Therefore calculating the average SRTA over the whole distribution hides the interesting signal contained in the data. Indeed, typically, the average SRTA would be negative. Given that observing values of the SRTA greater than zero is much less common than the opposite, the interesting information that the data provides with respect to comparison across groups is not provided by the mean. Rather, what really matters is how large the difference between the right tails of the distribution for the two groups is. Comparing the SPEC index across groups provides this information. Another popular choice in the literature is to calculate the SRTA for a given country as the aggregate of all of its firms. This approach is also unsatisfactory in the sense that the aggregate picture might be heavily influenced by a few large firms, washing away the information about comparative advantages or disadvantages of small firms. The SPEC index does not suffer from this problem either.

## 4.5 Findings

In the two following subsections, we present the findings that answer the two research questions raised in the introduction of this paper: *(i) In which life-cycle stages new innovators have a comparative technological advantage over incumbents? (ii) Are there significant differences in the comparative technological advantage of new innovators from different countries?* Before introducing the answers to these questions, we first describe the distribution of new and incumbent innovators in the NMPs sample. Table 12 reports the number of firms by geographic origin and type (new or incumbent innovators) across the five periods under consideration. To answer our two research questions, we merge the first and second component of the NMPs in the last period together, as explained in Sections 4.1 and 4.2.

**Table 12: Number of firms by geographic origin and category**

| <b>All firms</b> | <b>1981-1985</b> | <b>1986-1990</b> | <b>1991-1995</b> | <b>1996-2000</b> | <b>2001-2006 (1st+2nd)</b> | <b>Total</b> |
|------------------|------------------|------------------|------------------|------------------|----------------------------|--------------|
| US               | 61               | 92               | 62               | 75               | 80                         | 370          |
| JP               | 24               | 32               | 28               | 47               | 29                         | 160          |
| KR               | 0                | 2                | 5                | 7                | 5                          | 19           |
| TW               | 0                | 1                | 6                | 17               | 15                         | 39           |
| SG               | 0                | 0                | 1                | 4                | 3                          | 8            |
| KR/TW/SG         | 0                | 3                | 12               | 28               | 23                         | 66           |
| Total            | 85               | 127              | 102              | 150              | 132                        | 596          |

| <b>New Innovators</b> | <b>1981-1985</b> | <b>1986-1990</b> | <b>1991-1995</b> | <b>1996-2000</b> | <b>2001-2006 (1st+2nd)</b> | <b>Total</b> |
|-----------------------|------------------|------------------|------------------|------------------|----------------------------|--------------|
| US                    | 35               | 50               | 20               | 40               | 48                         | 193          |
| JP                    | 13               | 18               | 10               | 25               | 8                          | 74           |
| KR                    |                  | 2                | 3                | 2                | 3                          | 10           |
| TW                    |                  | 1                | 5                | 11               | 7                          | 24           |
| SG                    |                  |                  | 1                | 2                | 1                          | 4            |
| KR/TW/SG              | 0                | 3                | 9                | 15               | 11                         | 38           |
| Total                 | 48               | 71               | 39               | 80               | 67                         | 305          |

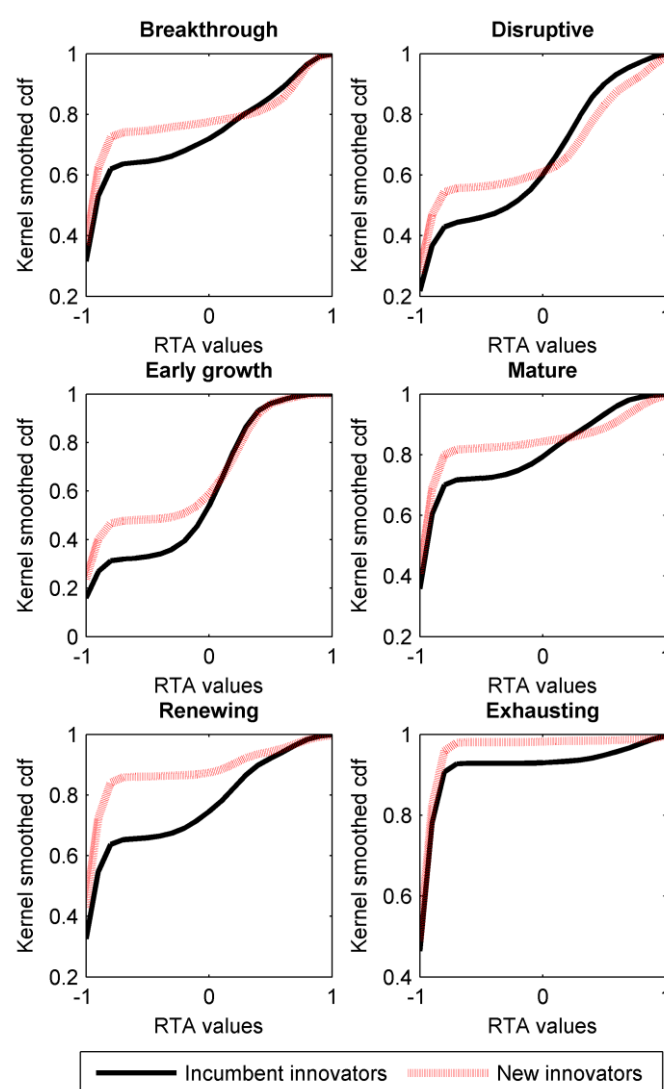
| <b>Incumbents</b> | <b>1981-1985</b> | <b>1986-1990</b> | <b>1991-1995</b> | <b>1996-2000</b> | <b>2001-2006 (1st+2nd)</b> | <b>Total</b> |
|-------------------|------------------|------------------|------------------|------------------|----------------------------|--------------|
| US                | 26               | 42               | 42               | 35               | 32                         | 177          |
| JP                | 11               | 14               | 18               | 22               | 21                         | 86           |
| KR                | 0                | 0                | 2                | 5                | 2                          | 9            |
| TW                | 0                | 0                | 1                | 6                | 8                          | 15           |
| SG                | 0                | 0                | 0                | 2                | 2                          | 4            |
| KR/TW/SG          | 0                | 0                | 3                | 13               | 12                         | 28           |
| Total             | 37               | 56               | 63               | 70               | 65                         | 291          |

#### 4.5.1 *New and incumbent innovators' revealed technological advantage*

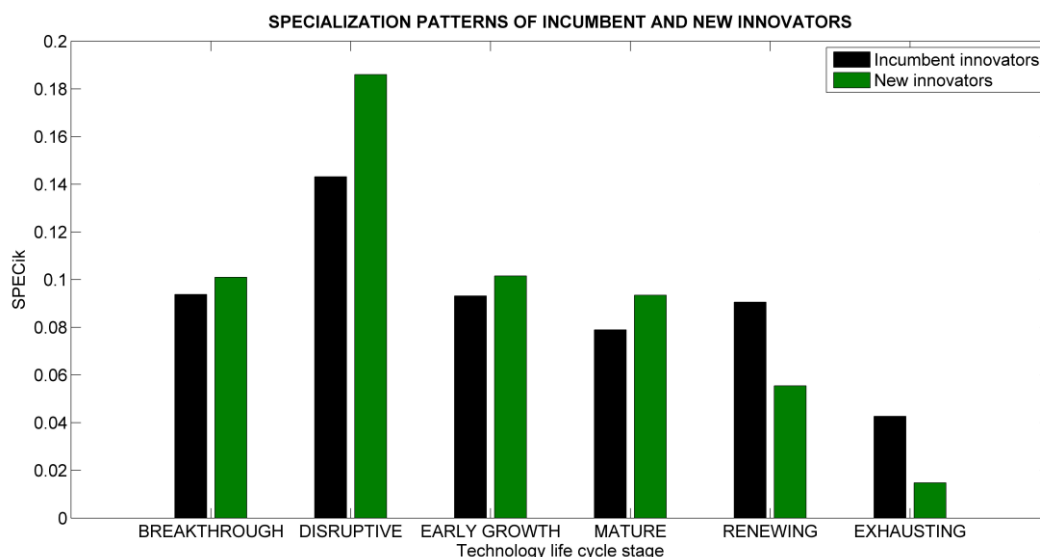
In order to have a reliable estimation for the distribution of SRTAs for new and incumbent innovators we initially plot all five periods together. This returns 305 observations for the new innovators and 291 for the incumbents. Figure 25 shows the kernel smoothed cumulative distribution functions for the two categories of firms. The vertical axis reports the probability to observe, across the whole sample, values of the SRTA smaller or equal than those reported on the horizontal axis. Therefore if one distribution is “smaller”<sup>17</sup> than the other for positive values of the SRTA it means that the former shows a comparatively stronger specialization pattern in the

<sup>17</sup> The correct terms would be *first order stochastic dominance* if one distribution were always below the other one and *second order stochastic dominance* if the two distributions cross at some point, meaning that one distribution is below the other only for values greater than a certain threshold. Stochastic dominance refers to the difference in probabilities to observe values of a given amount. If the distribution for one category is stochastically dominated (i.e. it falls below the other) for the whole or part of the range it means that the probability to observe large (small) values of the variable is higher (smaller) than for the other category.

given technology life cycle stage than the latter, as the probability to observe large SRTA values is higher. A first look at the figure reveals that the shape of the distributions changes across the different life cycle stages. However, in at least three cases, breakthrough, early growth and mature areas, the right tail of the distribution for both groups behaves quite similarly. The difference appears to be stronger in disruptive, renewing and exhausting areas. We test whether the behaviour of the two populations is statistically different by mean of the Anderson-Darling non-parametric two-sample test. The table of result is reported in Appendix A.4.4. The test confirms that the distribution of SRTA for new and incumbent innovators is statistically different for all the life-cycle stages except for the exhausting one. New innovators seem to have a comparative advantage in disruptive areas (as predicted by Christensen), whereas incumbents seem to be comparatively stronger, for mild levels of the SRTA, in renewing and exhausting areas, in line with industry life-cycle theory. A clearer picture of these differences is shown in Figure 26, where we plot the SPEC index for new and incumbent innovators.



**Figure 25: Estimated cumulative distribution functions for new and incumbent innovators**

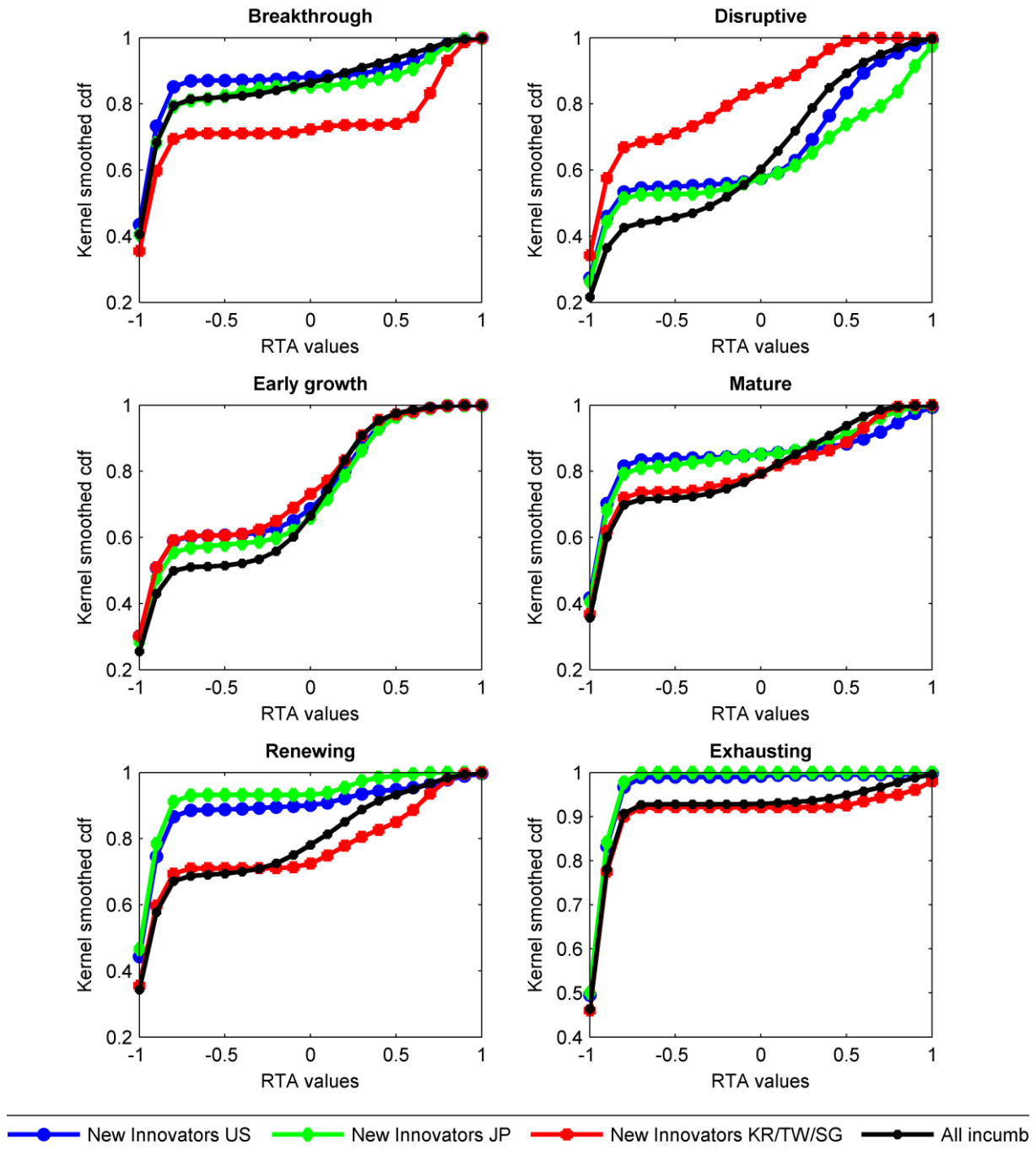


**Figure 26: Micro-founded specialization index for new and incumbent innovators**

Our micro-founded specialization index confirms what we inferred from the visual inspection of the cumulative distributions. New innovators have a greater probability than incumbents to have a comparative advantage in all life-cycle stages up to maturity. These differences are all statistically significant based on the Anderson-Darling test. However only for disruptive domains the comparative advantage is considerably strong. For renewing and exhausting domains the opposite is true and the comparative advantage is hold by incumbents. Yet the difference is significant only for the former. Therefore, if we only distinguish firms based on whether they are new or incumbent innovators, without considering their country of origin, the semiconductor industry follow a recommended specialization pattern which is consistent with industry life-cycle theory, Christensen's notion of disruptive technologies and Levinthal and March's definition of incumbents' myopia (Christensen, 1997; Levinthal and March, 1993). Indeed our findings are consistent with the theoretical prediction that new innovators perform comparatively better in technology domains in the initial stages of their life-cycle because incumbents are more likely to face learning traps that make them reluctant to explore new approaches to problem-solving. Our findings show that this is in general true but the comparative advantage is particularly strong only for disruptive domains. This answers our first research question. To tackle the second one we need to further distinguish firms based on their geographical origin. This is done in the next sub-section.

#### 4.5.2 Countries' revealed technological advantage

In Figure 27 we split new entrant innovators by geographical origin. Once again, in order to have enough observations for the estimation of the cumulative distribution function we plot all periods together (this constraint will be removed in the last part of the analysis). Furthermore, for the same reason, we need to group latecomer new innovators from Korea, Taiwan and Singapore into a single geographical area. This approach allows revealing the comparative technological advantages of new innovators from catching-up (i.e. Korea, Taiwan and Singapore), early entrant (i.e. Japan) and leader (i.e. US) countries. For the sake of further comparison, we also plot the distribution of SRTA for incumbent innovators. This distribution is the same shown in Figure 25.

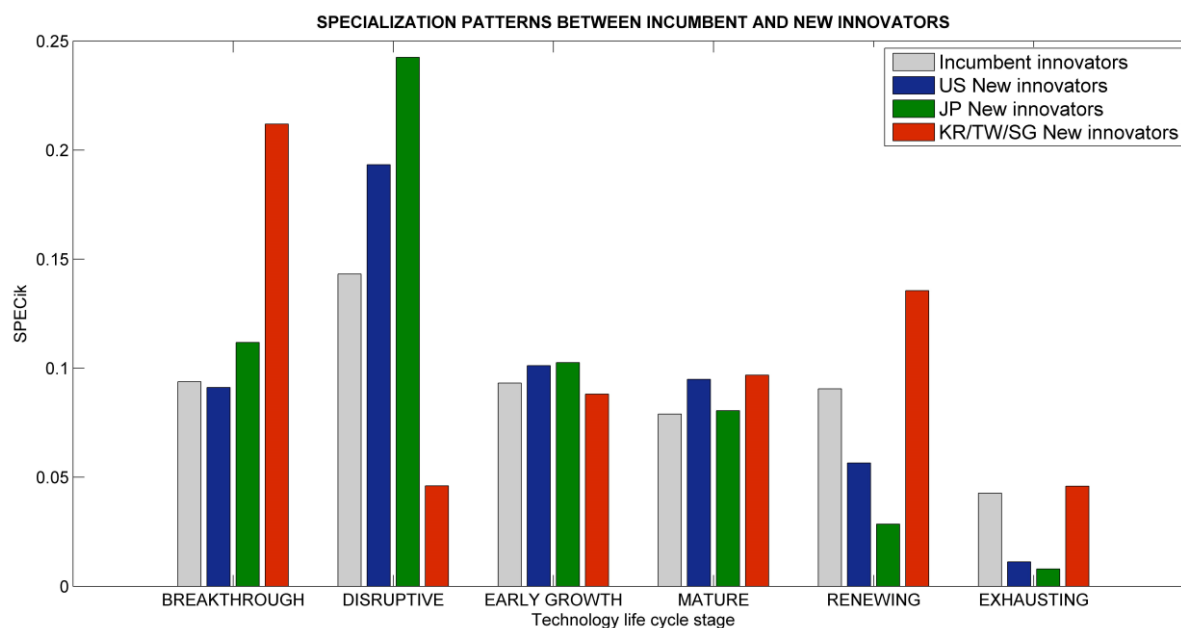


**Figure 27: Estimated cumulative distribution functions for new innovators from the US, Japan, and the three Asian tigers**

US and Japanese new innovators follow the same pattern of comparative advantage. The kernel estimated cumulative distributions of the SRTA values for US's and Japan's new innovators are extremely close in all the life cycle stages with the exception of disruptive areas. To the contrary, there is a remarkable difference between the distributions of the three Asian tigers and those of US and Japan, especially at the extreme stages of the life cycle. In breakthrough, renewing and exhausting areas, the distribution of SRTA values for Korean, Taiwanese and Singaporean new

innovators is always stochastically dominated by the distribution for US and Japanese new innovators. This means that Asian tiger's new innovators are comparatively more specialized in those areas that US and Japanese ones. The opposite is true for disruptive areas, whereas there is not much difference for early growth and mature ones. It is also interesting to compare specialization patterns between new innovators, now split by geographical origin, and incumbents. In technology domains in the early stages of their life-cycle, US and Japanese new innovators' specialization patterns closely follow incumbent innovators' one. On the other hand, for domains in the late stages (mature, renewing and exhausting), incumbents' distribution of SRTA values resembles more to the specialization patterns of new innovators from the three Asian tigers. This suggests that incumbent strategies are imitated more strongly by US and Japanese new innovators when it comes to specializing in emerging technologies, whereas they are followed more closely by Asian tigers' firms when the decision is about specializing in relatively older technologies.

As done in the previous section, to give a more precise answer to our second research question we look at the micro-founded specialization index for new innovators by geographical origin. This is reported in Figure 28. Once again, differences in the distributions plotted in Figure 27, which implies differences across SPEC indices, have been tested for statistical significance using the Anderson-Darling test (Appendix A.4.4).

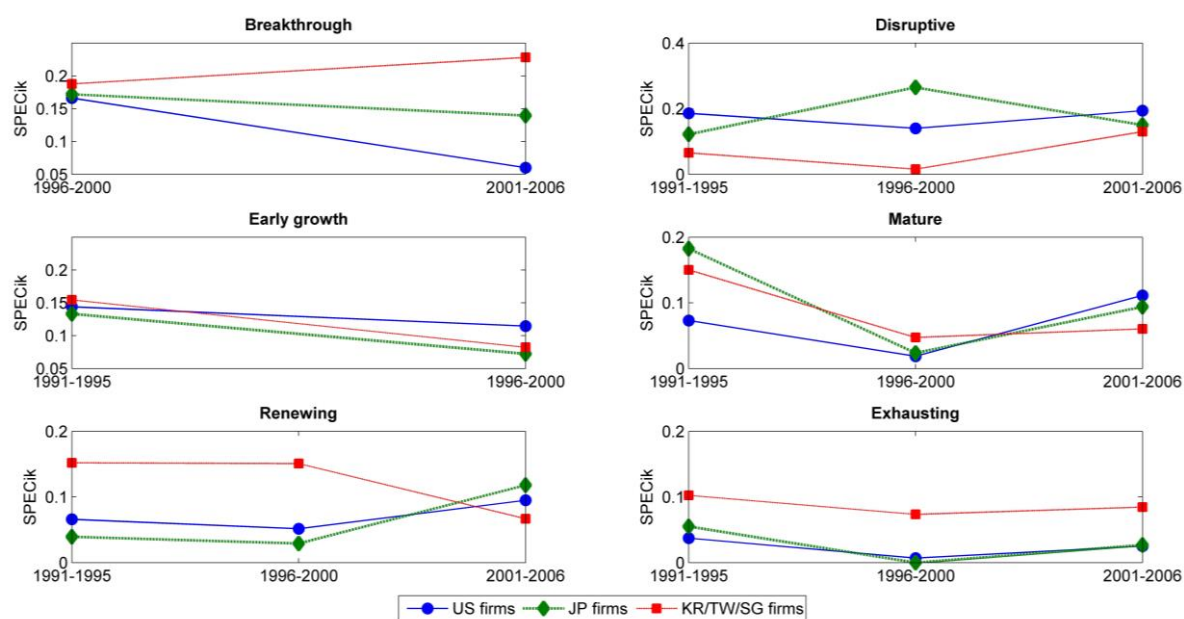


**Figure 28: Micro-founded specialization index for incumbents and new innovators by geographic origin**

Let us first consider breakthrough, renewing and exhausting domains. If we pick a firm at random out of each of the samples of new innovators, there is a larger probability that the randomly selected firm has a strong comparative advantage in those areas if we sample it from the Asian tiger group rather than the US or Japanese ones. Yet differences across the related distributions are statistically significant only for renewing domains. They are close to be significant in breakthrough and exhausting domains, when we compare Asian tigers' new innovators against US ones for the former and against Japanese new innovators for the latter. They are not significant in when comparing Asian tiger's and Japanese new innovators in breakthrough and Asian tiger's and American new innovators in exhausting domains. When we look at disruptive areas, the pattern reverses. Japanese and US new innovators enjoy a strong comparative advantage, whereas

Asian tiger's ones have a clear disadvantage. Yet, the Anderson-Darling test reveals that the advantage over the Asian tigers is significant only for Japanese new innovators, albeit close to significance for American ones. Differences are very mild for early growth and mature areas, although statistically significant in the case of early development domains for Japan. The advantage enjoyed by Asian tiger's new innovators over US and Japanese ones is consistent with the anecdotic knowledge of the development of the Semiconductor industry in these countries. As shown by Mathews and Cho (1999), Chang et al. (1994) and Cho et al. (1998), the strategy adopted by firms from Taiwan and Korea consisted in accessing relatively obsolete foreign technologies and reverse-engineer them to start their learning path. To the contrary, their comparative advantage in breakthrough domains, although not significant, deserves more attention. In particular, from the point of view of catching-up and knowledge upgrading, it is interesting to know when this advantage started to emerge.

Thus far, we provided a static analysis, due to the lack of a sufficient number of observations to have period-by-period reliable estimations for the new innovators. We can overcome this constraint by looking at all firms together, regardless of whether they are new or incumbent innovators. This way we are able to show a dynamic picture of micro-founded specialization patterns at the country level. Figure 29 shows the trend of the SPEC index over time across geographic areas.



**Figure 29: The evolution of the micro-founded specialization index over time**

A dynamic look at specialization patterns reveals that the comparative strength of Asian tigers in breakthrough domains is recent and started in the 2000s. Up to the end of the 1990s, firms from Korea, Taiwan and Singapore, were comparatively more specialized in renewing and exhausting domains. Interestingly, an increase of the SPEC index for these firms can also be observed in the 2000s for disruptive areas. What is also striking is that US and Japanese firms' comparative technological advantage in breakthrough areas (and disruptive ones, for Japan only) is decreasing in the 2000s in favour of areas at later stages of their life cycle (mature, renewing and exhausting). More information on the technical nature of each of the technology domains identified in the 2001-2006 period is included in the Appendix A.4.5.



These results shed light on the different strategies followed by the mayor players of the semiconductor industry. New entrants from emerging countries successfully catch-up with the leaders by initially specializing in renewing and exhausting technology domains. These areas of engineering research were left free by US and Japanese firms, which, up to the mid-1990s, were comparatively more specialized in disruptive and early growth areas. However, in the 2000s latecomer countries began to develop a distinct specialization in breakthrough areas and also an increasing focus on disruptive ones, though maintaining a comparative technological advantage in exhausting areas. A closer look at the data reveals that the large values of the SPEC index for Taiwan, Korea and Singapore in breakthrough and disruptive areas in the 2000s, is mainly due to their specialization in emerging areas belonging to the second component, rather than the first one. This highlights their ability to anticipate a possible radical change in the trajectory (in favour of semiconductor applications for devices such as e-readers, tablets, LCD monitors) and testify the effort they devoted to build capabilities in the new frontier thin-film transistor LCD technologies (Hung, 2006; Chang, 2005). The combination of these findings with what emerged from the analysis in Chapter 3, describe a clear picture of the learning strategies followed by latecomers in the Semiconductor industry. As shown in Chapter 3, up to the end of the 1990s, firms from Taiwan, Korean and Singapore were primarily focused on following well-established approaches to tackle central engineering problems in the semiconductors. However, the findings of the life-cycle analysis showed that, at the same time they were trying to renew these relatively older domains by mixing well-known approaches to problem solving with new ideas. These allowed latecomer firms building strong technological capabilities that quickly shift their comparative advantage to breakthrough areas in the early 2000s. This is confirmed by the ranking-changing strategies followed by some of these firms in the first half of the 2000s, as emerged from Chapter 3. Therefore, we can conclude that successful technological catching-up by firms from latecomer countries took a form that combined what Lee (2013) and Lee and Lim (2001) called path-creating and stage-skipping strategy. By focusing on renewing established engineering trajectories, they build sufficient technological capabilities to explore new ones. In contrast, for players from leading or early entrant countries (US and Japan), comparative advantage patterns reflect Klepper's industry life cycle theory. Entry focuses on emerging technologies, with a stronger advantage in disruptive domains, as predicted by Christensen (1997). For the sake of keeping the analysis concise, we did not show details on comparative advantage for individual firms. The interested reader can find a series of tables reporting SRTA indexes calculated for the mayor firms in the industry in the Appendix A.4.6.

## 4.6 Discussion and conclusions

Catching-up and leapfrogging in high-tech industries strongly depends on the direction of technological change and on the emergence of new technology domains and decline of old ones. In fast changing technical and business landscapes today's capabilities do not necessarily ensure long-run survival. This highlights the importance of studying the relationship between technology life cycle and the dynamic of comparative advantage patterns of new and incumbent innovators. Our study is one of the few empirical contributions, together with Lee (2013), to the discussion of technology life cycles at the domain level. Patent citation networks offer a fertile ground for such analysis. We theoretically defined the life-cycle of technology domains and its relation with product and industry life-cycles. Furthermore, we built a methodology to identify technology domains and trace their life-cycle by means of disentangling the complexity of large patent

citation networks. This provided new insights on the dynamics of comparative advantage in the semiconductor industry.

First, we confirmed the empirical validity of entry and comparative advantage predictions from the theories of industry life-cycle and disruptive technologies. Second, we showed that, until the end of the 1990s, US and Japanese firms were comparatively better in emerging technology domains, whereas firms from Taiwan, Korea and Singapore, tended to specialize in relatively older domains, mainly in their mature, renewing and exhausting stages. These comparative advantage patterns changed strongly in the beginning of the 2000s, when firms from the three Asian tigers, next to their advantage in declining domains they also started developing a comparative advantage in emerging ones. This proves that latecomer firms from these countries have engaged in a mix between path-creating and stage-skipping catching-up, as theorized by Lee and Lim (2001). These results are also in accordance with the empirical analysis of technology cycle time and catching-up made by Lee (2013) in which the author shows that the successful catching-up of Korea and Taiwan built on upgrading the specialization pattern from older to newer technologies, exploiting short-life cycles. Our findings are also in line with the description of how Korean and Taiwanese firms managed to build their technological capabilities, as discussed by Chang et al. (1994), Mathews and Cho (1999), Cho et al. (1998), Chang and Tsai, 2002, Bell and Juma (2008) and Hobday (2000). These authors agree in highlighting the instrumental role played by Korean and Taiwanese firms' early specialization in old foreign licensed technologies to develop internal R&D capabilities lately used to upgrade their specialization. The Asian tigers' relatively strong position in domains that were emerging in the early 2000s, testifies their ability to be forward-looking.

Yet, it is important to notice that in this work we did not assess the future impact of emerging domains. Our goal was to analyse whether new entrants' comparative advantage in those domains significantly differs from incumbents' one. It is needless to mention that emerging technologies are intrinsically risky and there is no guarantee that their development will be sustained in the future. A detailed analysis of how emerging areas affect the future direction of the technological trajectories goes beyond the scope of this paper. However, a preliminary analysis, that was not reported here, revealed that some areas did generate sustained new trajectories whereas others failed to do so. Since this has crucial implication for catching-up, a full analysis of the knowledge interaction between technology domains and the transferability of capabilities between areas is an open question for future research.

Finally, we want to praise the strength of using interdisciplinary approaches to disentangle today's technological and economic complexity. Several tools have been developed for this purpose, mainly at the intersection of economics with mathematics, physics and network science. The application of economic thinking to a combination of these tools, the community detection technique and the network of main paths, proved to be extremely insightful to analyse an economic question that occupied scholars at least since Vernon's seminal work (1966), namely the one of the relationship between life cycles and comparative advantage. The correspondence of our findings with the extensive anecdotal knowledge of catching-up in the semiconductor industry contributes to validate our methodology to trace the life-cycle of technology domains and make a case for its use to study the technology dynamics of other high-tech industries or apply it at a wider scale to the question of the co-evolution of technologies.

# 5 DO I FIT? INNOVATIVE SURVIVAL, TECHNOLOGY DYNAMICS AND THE EVOLUTION OF KNOWLEDGE

We study firms' survival as core innovators (i.e. innovative survival) in the semiconductor industry as a function of knowledge diversification and accumulation. The semiconductor industry is characterized by high level of technical modularity, which is reflected in poor knowledge proximity across technology domains. We hypothesize that, in such context, diversification and knowledge accumulation provide better chances for innovative survival. Moreover, we argue that, when knowledge is hardly portable across domains, the spread of survival chances between diversified and specialized players should be larger when technological change is knowledge replacing. We contribute new measures of knowledge persistence, breadth, depth and modularity, which take into account the topological structure of the system of technology domains. We do that by combining the Network of Main Paths (NMPs) analytical tool with the genetic approach (GA) to patent citation networks defined by Martinelli and Nomaler (2014). Our findings show that diversification significantly improves core innovators' odds of innovative survival on the main paths of citations but only for large levels of knowledge breadth. The effect is conditional to the level of knowledge persistence. In contrast, firm's knowledge depth alone, as well as the size of its technical knowledge base and the business method adopted, does not affect innovative survival. However, there are positive complementarities between knowledge breadth and depth. Important policy and managerial implications of these findings are discussed.

## 5.1 Introduction

Technology, innovation and capability upgrading are widely believed to be among the main drivers of economic growth (Kim and Nelson, 2000). Several studies have analysed the empirical relationship between innovation and performance. In some cases the simplest measure of performance has been used, i.e. survival. A variety of factors has been proven to affect firm survival. A not-exhaustive list includes product innovation (Fontana and Nesta, 2009, Cefis and Marsili, 2005), new technology development (Levitas et al., 2006) product differentiation (Cottrell and Nault, 2004), capabilities (Sapienza et al., 2006), pre-entry experience (Bayus and Agarwal, 2007, Buenstorf, 2007) and industry life cycles (Agarwal and Gort, 2002). From an evolutionary perspective, survival is seen as the ability to adapt to change (or successfully introduce mutations) and fit the new environment. The urge to adapt depends on the scale and direction of change. Audretsch, (1995) showed that a highly innovative environment exerts a disparate effect on the post-entry performance of new entrants. New firms that are able to adjust and offer a viable product experience higher rates of growth and a greater likelihood of survival. Therefore, the survival of firms in the competitive environment is substantially affected by the evolution of the technology underlying their products and the related manufacturing process. (Suarez and Utterback, 1995).

The lesson that we can learn from the studies on firm survival is that the ability to innovate and to adjust to changes in the technological environment are key explanations of market survival. Consequently, understanding the determinants of innovative survival is of crucial importance. We characterize innovative survival as firm's persistent presence in the set of core innovators. By the latter, we mean those players that are capable of affecting the direction of technological change by consistently produce influent technological improvements. In other words, core innovators define the approaches followed to solve engineering problems. To identify influent technical improvements we select, out of all semiconductor-related patents, those that are centrally located on the paths of technical development of semiconductor technologies. This is done using a bibliometric approach called Network of Main Paths (NMPs). We estimate the probability of firms' innovative survival as a function of their technical knowledge breadth, depth, experience and business type. We also assess how the scale of knowledge replacement along the main paths of technological improvement, affects innovative survival. We use patent data from the second version of the NBER patent citation database (Hall et al., 2001), which cover the period between 1976 and 2006. Since US are a crucial market for semiconductors we assume that any technologically or economically important invention in this field is patented at the USPTO.

Surprisingly, the number of studies that focus on innovative survival as dependent variable is considerably smaller than those focusing on firm market survival, even though the former has been hypothesized to explain the latter. As Lee and Lim (2001) pointed out, a durable and sustainable market success crucially depends on the ability to upgrade firm's technological capabilities constantly. Malerba and Orsenigo (1999) showed that innovative turbulence is an important composite phenomenon, in which innovative entrants/exiters and lateral entrants/exiters play different roles. Most of the entrants are occasional innovators, while persistent innovators are only a few but have the largest patent shares. That is why they are referred to as core innovators. Explaining innovative survival (sometimes also called innovation persistence) is not an easy task. Studies on the topic have found mixed evidence (Antonelli et al., 2012; Roper and Hewitt-Dundas, 2008; Raymond et al., 2006; Malerba et al., 1997). Most of them measured persistence using self-reported innovation outcomes or indirect measures of innovation that are

more closely related to strategic decisions, like R&D investments, or affected by large noise, like patent counts. For instance, Antonelli et al. (2012), measured product and process innovation persistence using a self-reported binary variable from a survey of firms. They found that the higher level of persistence is found for R&D-based innovation activities. Roper and Hewitt-Dundas (2008), using data from the European Community Innovation Survey (CIS) found that product and process innovation exhibit strong persistence at the firm level, but they find that persistence is not significantly stronger among highly active innovator. On the other hand, Raymond et al. (2006), also using the CIS, found no evidence of true persistence in firms' ability to achieve technological product or process innovation. This seemingly contradictory evidence might be explained by hypothesizing that, if technical change is knowledge-replacing, the cumulative advantage of incumbent core innovators is broken. Therefore, in this scenario, we expect no differences in the ability to innovate between new entrants and incumbents as they both start at the same level of knowledge. The need to further understand the determinants of innovative survival provides the motivation of this paper.

We focus on the global semiconductor industry as a case study. Several factors threaten semiconductor firms' ability to persistently generate influential technical solutions. Firms need to keep pace with an exponential rate of improvement (Koh and Magee, 2006) fuelled by high competition and short product life cycle (Brown and Linden, 2009). Furthermore, in the last two decades, the industry has experienced an unprecedented level of technology modularization that spurred functional specialization along the value chain and entrance of specialized suppliers. This further increased competition. There are four main types of players in the industry, *Integrated Device Manufacturers* (IDM), which design, manufacture and commercialize their own chips, *fabless* companies, which specialize in the design of semiconductor devices, *foundries*, which manufacture them on behalf of third parties and *specialized suppliers* of equipment and materials. The technology modularization has fostered the entrance of fabless, foundries and specialized suppliers by reducing the economies of scope due to the decreased interdependence of different technological components (Adams et al., 2013; Langlois and Steinmueller, 1999). Modularity has also increased the variety of possible applications for semiconductors. The diffusion of Complementary Metal Oxide Semiconductor (CMOS) production processes reduced physical and cognitive interdependence between product design and manufacturing. This was caused by the standardization of interfaces between components. The introduction of systems-on-a-chip (SoC) and Application Specific Integrated Circuits (ASICs) furthered increased the market segmentation and fostered product specialization (Adams et al., 2013; Brown and Linden, 2009).

However, this positive view on modularity and specialization is contrasted by more cautious theories in the literature on production networks. Ernst (2005b) argues that: "*competitive dynamics and cognitive complexity create modularity limits. [...] interfirm collaboration requires more (not less) coordination through corporate management, if codification does not reduce complexity—which it fails to do when technologies keep changing fast and unpredictably*". At the technological level, modularity means that technology is developed into separate and independent components. Therefore mastering the design of one component requires little knowledge about the internal functioning and the design process of the others. Even though this clearly fosters specialization, it also means that knowledge is highly domain-specific. Therefore, if a technological shock affects a given component, innovative survival of firms specialized in that component requires being able to quickly jump to the new generation of the domain-specific technology. This is because what had been learned so far is not useful to innovate in other domains. In other words, there are no escape routes to related knowledge domains. When between-technology knowledge flows are limited, the ability to adopt a new generation of the domain-specific tech-

nology, for innovators specialized in that domain, crucially depends on the intertemporal persistence of domain-specific knowledge. If knowledge persists across time the urge of diversification is smaller because the knowledge required to master the new technology is did not change much compared to the old one. This is the case when technological change is cumulative.

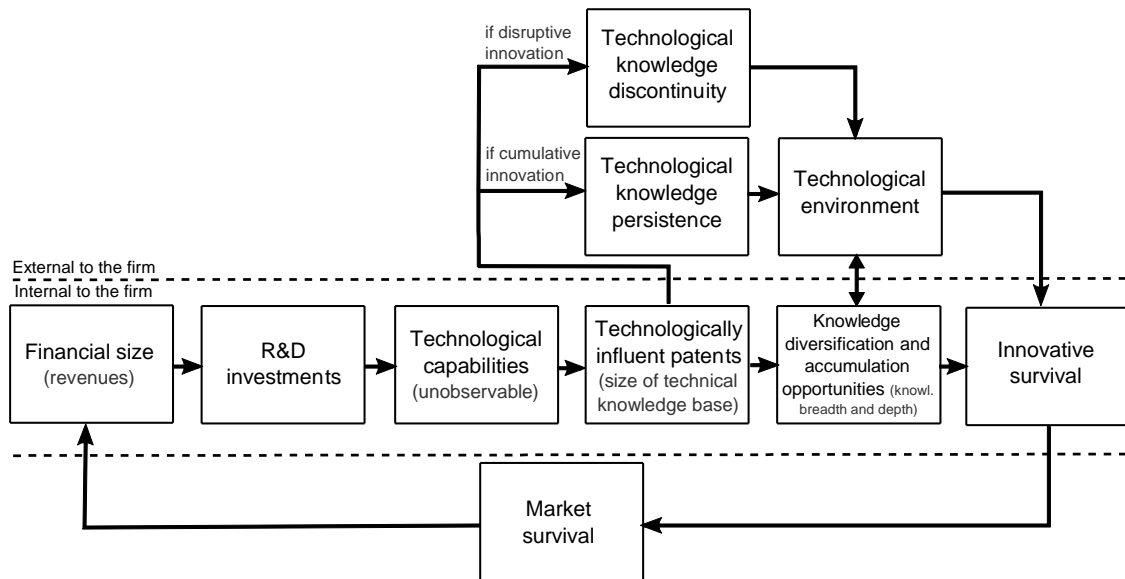
This paper focuses on the relationship between modularity, knowledge persistence and innovative survival. The research question that we seek to answer is whether technical knowledge diversification (i.e. knowledge breadth) affects firms' innovative survival, in particular in periods of large technological turbulence. We also argue that diversification is only one dimension of the knowledge space. Knowledge accumulation (i.e. knowledge depth) provides a complementary dimension. In other words, if we think about knowledge as a tree, the breadth dimension corresponds to the number of branches the tree has. The depth dimension tells us how deep the roots feeding each of the branches are. The relationship between knowledge depth and survival is more ambivalent. Knowledge accumulation can provide a shield against technological turbulence by allowing early detection of decline of a given technology domain due to the high level of experience. However, the opposite could be true if the former investments to accumulate that knowledge generate resistance to change and learning myopia (Levinthal and March, 1993). We contribute to the debate on innovative survival by defining a new method to identify core innovators that goes beyond simple patent counts or self-reported outcomes. Furthermore, we develop a set of indicators to measure knowledge modularity and persistence at the patent level, and firms' knowledge breadth and depth. These measures are based on the genetic approach (GA) to patent citation networks developed by Martinelli and Nomaler (2014). The GA takes into account the topography of the system of technical knowledge. It is used here to map relationship across technology domains and the structure of firms' knowledge base. To identify the set of technology domains within the Semiconductor Industry we use a network community detection technique developed by Newman (2004). The paper is structured as follow. Section 5.2 clarifies the theoretical framework and defines the hypotheses that guide our analysis. Section 5.3 present the data used. The set of models that we estimate are introduced in Section 5.4. In Section 5.5 we explain how we measured the model variables. Trends of knowledge modularity and persistence in the semiconductor industry are discussed in Section 5.6. Finally, in Section 5.7 and 5.8 we introduce and elaborate the preliminary empirical analysis and the findings from the regressions.

## 5.2 Theoretical framework and research hypotheses

The theoretical framework that we apply in this paper is based on seven key concepts and their interaction: *engineering trajectory*, *knowledge persistence*, *technology domains*, *technical knowledge modularity*, *firms' knowledge breadth and depth* and *innovative survival*. We first theoretically define each of these concepts and then discuss their interactions. By engineering trajectory, we mean the main paths of technical improvements followed in an industry to tackle product and process design challenges. It represents how innovators solved technical problems related to create and manufacture better performing products, i.e. semiconductor devices in our case. Engineering trajectories are, therefore the projection in the design space of technological trajectories defined by functional and technical trade-offs that emerge at the product level. As Dosi (1982) theorized, a technological trajectory is defined as "*the 'normal' problem solving activity determined by a [technological] paradigm. [It] can be represented by the movement of multi-dimensional trade-offs among the technological variables which the paradigm defines as relevant*". The set of product features one wishes to improve and the engineering problems that need to be

addressed to achieve those improvements are deeply intertwined and affect each other in a complex fashion. The historical paths of technical approaches used to solve those design challenges are what we define as engineering trajectories. When current approaches to problem solving cumulatively build on previous ones, knowledge developed in the past is still useful in the present. In this case, we can state that there is a strong persistence in technical knowledge. Therefore, this instance of technical change can be described as knowledge reinforcing. In contrast, when innovators explore new approaches to solve engineering problems, which do not build on previously well-established paths, the trajectory of improvements experiences a discontinuity. Previous knowledge becomes less useful to solve current problems. This pattern of technical change can be said to be knowledge-replacing. Ultimately, the presence or lack of changes in the approaches to problem solving depends on the strategic decisions of the players involved. However, some firm are more influent than others in shaping the direction of the engineering trajectories. We measure firms' technical influence as the ability to have central patents on the paths of technical improvements. Patents' monetary and technological value is notoriously highly skewed (Gambardella et al., 2008; Silverberg and Verspagen, 2007; Hall et al., 2005; Reitzig, 2003). Some technologies attract much more innovative effort than others do and some are able to set the path of current and future engineering solutions to selected technological problems. Firms' ability to persistently having influent patents is what we define as innovative survival. When we think about the design space of an industry, we intend that the set of products and related engineering challenges is heterogeneous. In particular, the set of applications of semiconductor devices is highly diversified, including products as diverse as microprocessors, memories, light-emitting diodes (LED), liquid-crystal displays (LCD), personal computers, smartphones, tablets, and application-specific integrated circuits (ASICs). Therefore, the set of engineering challenges and approaches to problem solving is very diverse. We define technology domains within an industry as areas of applied research that share a set of common technological problems that are tackled applying similar approaches. Technical modularity measures the extent to which knowledge is domain-specific. In a highly modular system, the technical knowledge that is needed to succeed as an inventor in a given domain is very unrelated to what is needed in other domains. The notion of a system of domains allows defining the concept of firms' knowledge breadth and depth. The former refers to how diversified a firm's technical knowledge is across domains. The latter indicates how much of the knowledge historically generated within a domain is known by the firm.

To discuss how the seven key concepts of our theoretical framework interact with each other, we need to define a broad picture of firm's inventive process and its relationship with firm survival, both as core innovator and in the marketplace. This picture is illustrated in Figure 30.



**Figure 30: Theoretical framework**

Large firm's size allows for conspicuous investments in R&D, which, depending on the level of technological capabilities, can generate high or low quality technology outputs. These can be measured by the level of technical influence of its patents. We argue that the ability to generate such influent technologies depends not only on R&D investments but also on the unobservable technological capabilities. The number of technologically influent patents granted to a given firm can be used as a proxy of the size of its technical knowledge base. The larger the knowledge base the more diversification and knowledge accumulation opportunities can be enjoyed. Depending on how the disruptive or cumulative nature of the firm's innovative output, technologically influent patents can either set a new engineering trajectory or reinforce the prevailing approach to problem solving. This creates a systemic shock or reinforces the known landscape of the technological environment. In turns, this can either create new opportunities for knowledge diversification or strengthen knowledge accumulation. This is a bidirectional relationship as firms' knowledge diversification and accumulation effort also changes or reinforces the existing technology landscape. Ultimately, innovators' resilience to technology shocks or probability to successfully introduce change, depend on diversification and the level of knowledge accumulation. Note that the shock is exogenous to firms that react to novelty but endogenous to the firm exploring the new trajectory. However, in the latter case success is not guaranteed. Therefore, diversification and knowledge accumulation can still be beneficial. The former, by sharing the innovator's bets across several technology domains. The latter, by increasing the likelihood that the new trajectory will eventually be the winning one. Finally, depending on exogenous market response to technology introduction and development, innovative survival leads to financial capital accumulation, which injects additional resources into the system.

The ease of exploiting diversification opportunities depends on the level of relatedness between firms' current products and technologies and those in which they wish to diversify (Teece et al., 1994; Breschi et al., 2003; Neffke and Henning, 2013). Depending on the structure of the product and technology spaces, diversification may not be trivial. One might wonder then what makes technical knowledge diversification beneficial. Diversification is usually intended in the literature as the process of adding new products to the firm's portfolio or entering new sectors of the economy (Frenken and Boschma, 2007; Saviotti and Pyka, 2004). However, the commercialization of new products or entry in new markets does not imply that the firm is diversifying.



This could also be achieved by shifting firm's specialization. Moreover empirical evidence shows that novelty is often introduced by new ventures rather than incumbents especially at the beginning of the technology life cycle, as argued by Utterback and Abernathy (1975) and Christensen (1993) and empirically showed, for instance, by Klepper (2002 and 1997), Rosenbloom and Christensen (1994), and in Chapter 3 of this thesis. We argue that the peculiar advantage of technical knowledge diversification is the increase in the probability to persist fitting the environment for a simple economic principle of risk sharing. This ultimately creates dynamic returns as survival allows for knowledge and physical capital accumulation that sustain the growth path. This argument moves the spotlight from market survival and macroeconomic shocks to innovative persistence and technology shocks. Following this reasoning, we formulate our first hypothesis.

**Hypothesis 1:** Innovative survival is positively affected by knowledge diversification (i.e. large knowledge breadth).

Knowledge diversification measures how much firms can share risk across technology domains. Moreover, it also indicates the extent to which firms can enjoy knowledge complementarities across several technology domains and knowledge spillovers coming from advances in related domains. This is what Frenken et al. (2007) define as related-variety. From a theoretical point of view, complementarities and spillovers are possible only under two circumstances. First knowledge from different domains must be somehow related. This is not the case under modularity. Second, that the firm has a sufficient level of absorptive capacity to effectively internalize knowledge coming from outside the boundaries of the current firm's domain(s) of specialization. As explained by the pivotal work of Cohen and Levinthal (1990), the latter can only be achieved in presence of pre-existing internal learning efforts. In other words, firms need to have sufficient knowledge depth to enjoy knowledge spillovers from related domains. Large knowledge accumulation can allow early detection of decline of a given technology domain and provide the necessary absorptive capacity to eventually attempt to take alternative paths. On the other hand, former investments to build knowledge depth can generate resistance to change that is particularly high if the investment is seen as a sunk cost, i.e. when high modularity makes knowledge less portable across technical domains. It is therefore difficult to a priori-define a hypothesis about the relationship between knowledge depth and innovative survival. They might be characterized by a positive or negative relationship or the combined effects of the positive and negative aspects of knowledge depth might generate an inverted-U shape. We therefore formulate a general hypothesis, without specifying whether the relationship is positive or negative and defining its functional form.

**Hypothesis 2:** Knowledge accumulation (i.e. large knowledge depth) affects innovative survival.

Technical knowledge breadth and depth can also be interpreted as measurable outcomes of exploration and exploitation strategies respectively. In his seminal work March (1991) defined the organizational trade-off which is intrinsic in the decision to allocate resources to the "*exploration of new possibilities or the exploitation of old certainties*". What attracted particular interest in the organization learning literature is the hypothesis that ambidexterity, i.e. the ability to combine both exploitation and exploration, is the key to sustained performance (Raisch et al., 2009). A similar explanation focuses on dynamic capabilities, i.e. the ability to update and reconfigure firms' routines, knowledge and strategies as the main determinant of long-term survival. O'Reilly and Tushman (2008) proposed a synthesis of the two approaches that sees ambidexterity as a manifestation of dynamic capabilities. Following these theories we hypothesize that technical knowledge breadth and depth have strong complementarities. Obviously, it is not easy to simultaneously pursue both. The innovator's attempt to master new domains is likely to re-

duce, at least initially, its overall knowledge depth because a previously unknown area is explored. Yet, successfully breaking the trade-off and being able to enlarge both breadth and depth allows risk-sharing and enjoying economies of learning. He and Wong (2004) showed that the interaction between explorative and exploitative innovation strategies is positively related to sales growth rate. Similarly, achieving complementarities between technical knowledge diversification and accumulation should increase the odds of innovative survival. This argument leads to our third hypothesis.

**Hypothesis 3:** Innovative survival is positively affected by the ability to break the trade-off and achieve large values of both knowledge breadth and depth.

So far, we have discussed the effects of knowledge diversification and accumulation on innovative survival in isolation from technological change. Clearly, the probability of innovative survival crucially depends on how useful the current and past knowledge is to be successful in the future, as measured by knowledge persistence. As shown by Breschi et al. (1998) and Malerba and Orsenigo (2000) patterns of organization of innovative activities are determined by the nature and peculiarities of the underlying technology. Technological regimes are defined as a combination of some fundamental properties of technologies, namely opportunity conditions, appropriability conditions, degrees of cumulativeness of technological knowledge and characteristics of the knowledge base. Lee and Lim (2001) stress that as far as catching-up by latecomers is concerned, cumulativeness plays a particularly important role as it increases the predictability of the technological trajectory. We add that this is true for all firms, regardless if they are incumbents or latecomers. More generally, we hypothesize that high knowledge persistence favors current core innovators' ability to persist being influent. Obviously, the opposite is also true.

**Hypothesis 4:** Low knowledge persistence increases the hazard of innovative exit.

Furthermore, as argued by Malerba and Orsenigo (1993), the higher the degree of cumulativeness of technical change the larger the set of technology paths that current core innovators can take. However, it is important to distinguish between *ex-ante* and *ex-post* paths. As we argued in the introduction, high levels of knowledge modularity make it difficult to create escape paths to other knowledge domains *after* a shock in the technological trajectory has occurred. Therefore, to improve survival chances it is necessary to diversify *ex-ante*. Hence, we postulate our last hypothesis.

**Hypothesis 5:** Knowledge diversification provides a larger survival premium compared to specialization when technological change is knowledge-replacing (i.e. when knowledge persistence is low).

To test these hypotheses empirically, we first need to perform a number of tasks. We need to identify influent patents, core innovators and technology domains within the Semiconductor industry. Furthermore, we need to measure knowledge persistence, modularity and firms' knowledge breadth and depth. These tasks and the methodology used to tackle them are summarized in Table 13. We will discuss them in details in the next section.

**Table 13: Methodology roadmap**

| Task   | Methodology used                                  |
|--|---|
| <b>Data preparation</b>                                |   |
| Identify influential patents and core innovators       | Network of main paths (NMPs)                      |
| Identify technology domains                            | Community detection on the NMPs                   |
| <b>Measurement of the model variables</b>              |   |
| Measure technical knowledge modularity and persistence | Genetic approach to patent citation networks (GA) |
| Measure firms' technical knowledge breadth and depth   | GA  |

## 5.3 Data preparation and description

### 5.3.1 Identification of influential patents and core innovators

We use data from the NBER-USPTO database containing metadata of granted utility patents and their citations from 1976 to 2006 (Hall et al., 2001). Our initial dataset is composed of all organizations that have been granted at least one patent classified in one of the semiconductor US technological classes between 1976 and 2006<sup>18</sup>. This dataset is made of 4170 unique assignees, 114097 patents and 779076 citations for the entire period. This dataset is reduced to select only technologically influent patents and their assignees. We define technological influence as the ability to affect the direction of technological change. This is intended as the act of indicating which are the most promising approaches to address the key engineering challenges that affect a given industry. To identify technologically influent patents we make use of a methodology that analyses the *network of main paths* (NMPs). The NMPs' approach has been originally proposed by Hummon and Doreian (1989), further developed by Verspagen (2007) and lately applied by Fontana et al (2009), Martinelli (2008 and 2009) and Bekkers and Martinelli, (2010) to several empirical cases studies of technology dynamics in different industries. The NMPs identifies the routes through which knowledge diffuses in large citation networks (made of patents or publications). When applied to patent citation networks this methodology allows analysing the evolution of the main sequences of engineering improvements within the technology domain(s) underlying a given industry. The first building block of this approach relates to the meaning of patent citations. If patent B cites patent A then one of the claimed features of the invention described in the former improves upon the latter. In other words, as described in the USPTO Manual of Patent Examination Procedure, A represents the state-of-the-art concerning a claimed feature of the technology described in patent B, at the moment in which patent application B was filed (USPTO, 2014). Therefore, citations can be interpreted as a measure of technological relatedness and provide insights on the direction of technological change. Obviously, patents can cite and be cited by many patents. Hence, if we want to follow the main trajectories of technology evolution among a set of patents, we first need to decide which direction to take at every junction. This is what the NMPs algorithm does. First, we calculate the weight of every citation using the search path node pair (SPNP) algorithm, as developed by Batagelj (2003). The SPNP returns the number of times that each citation link lies on all possible paths connecting any node to any one else. This is easily calculated by multiplying the number of patents that reach (through direct and indirect citations) the cited patent by the number of patents that are reached (directly or indirectly) by the citing patent. Therefore, a high SPNP weight indicates that the given citation and the two patents involved are located in a highly connected and connecting area of the network. This means that the given citation has a strong technological influence, as many paths of technological improvement pass through it. The NMPs is identified by following the paths emanating from start nodes (nodes that are cited but not cited), taking at each junction the direction of the citation which carries the highest weight, till an end point (a node who cites but is not cit-

---

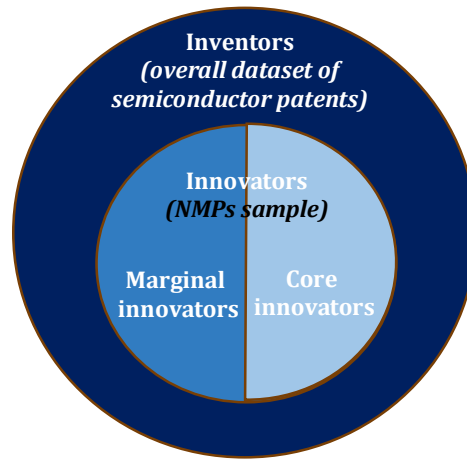
<sup>18</sup> The US patent classification distinguishes five classes related to semiconductor technologies: 257, 'Active solid-state devices (e.g. transistors, solid-state diodes)'; 438, 'Semiconductor device manufacturing: process'; 326, 'Electronic digital logic circuitry'; 505, 'Superconductor technology: apparatus, material, process'; 716, 'Design of semiconductor devices'.

ed) is reached. By repeating this procedure for each start point, we identify the NMPs. The NMPs can be made of separate network components when some of the main paths do not intersect. Engineering solutions described in patent belonging to different components can therefore be interpreted as unrelated. It is important to notice that components of the NMPs are not necessarily separated if we look at the original network, but the nodes that connect them have a negligible importance from the point of view of technological trajectories. We interpret patents that show up in the NMPs sample as the technologically influent one because of their importance in setting the engineering trajectory followed by innovators in the industry in the window of time under observation. We identify the NMPs for six left-cumulated periods: 1976-1980, 1976-1985, 1976-1990, 1976-1995, 1976-2000 and 1976-2006. Cumulating periods starting from the same initial grant year (i.e. 1976), allows assessing how newly granted patents affects the stability of the main paths of technical improvements. By doing so, we can identify influent patents granted in the last five years of each period. For each period, we only focus on the largest component of the NMPs and the second one when deemed important. This applies to periods 1991-1995 and 2001-2006, when, as shown in Chapters 3 and 4, the second component accounted for a significant share of the influent inventing activities. This is the sample of technologically influent patents that we use as the basis for identifying core innovators. From now on, we will refer to this sample as the *NMPs sample*. It is made of 673 unique assignees and 9555 unique patents granted in the six five-years periods mentioned above.

The set of 673 assignees of NMPs patents represents the group of *innovators* in the industry. In other words, a given firm might be engaged in patenting in one of the semiconductor classes but none of its patents might be important enough from the perspective of technological evolution in the industry to make it to the NMPs sample. In this case, the given firm is not considered an innovator (it might at best be considered an inventor). This builds on Schumpeter's distinction between invention and innovation based on their recognized usefulness. Innovators can be further distinguished between *core* and *marginal innovators*. Core innovators are entities that appear in the NMP sample for more than one period (i.e. over a time horizon of at least 10 years) or entities that show up in just one period (i.e. 5 years, except for the last period which is made of 6 years) but hold a significant share of patents in at least one technological domain<sup>19</sup>. All the other entities are marginal innovators whose patents are important enough to make it to the NMPs sample but whose inventive outcome in the semiconductor technology domain lacks either persistence or success. This is similar to the definition of core innovators from Malerba and Orsenigo (1995) and Malerba et al. (1997). The conceptual relationship between inventors, marginal and core innovators and the NMPs sample is graphically summarized by Figure 31.

---

<sup>19</sup> There are 142 unique core innovators between 1976 and 2006. Within this set, 129 show up in more than one period, although not necessarily in a row, whereas only 13 of them are only present once but have a significant share of patents in at least one technology domain. However, none of these 13 appears in the final subsample that we use for the econometric analysis. Therefore, the sample used for the regressions, which is limited to the window of time between 1980 and 2000, is only made by core innovators that show up in the NMP more than once.



**Figure 31: The relationship between inventors and innovators**

We can also categorize innovators in terms of their time of entrance. We distinguish between new, incumbent and exit innovators. New innovators are organizations that show up in the NMPs sample for the first time. They might later prove to be core or marginal innovators. Incumbent innovators are organizations that appeared in the list of assignees of the NMPs sample at least once in the previous periods. Exit innovators are organizations that ceased to be part of the NMPs sample in the period under observation.

Table 14 provides detailed information on the number of technologically influent patents (i.e. the size of the NMPs sample), their share of the total number of semiconductor patents, the number of innovators and their sub-categories for each of the five periods under observation. Two major trends emerge from Table 14. First, the number of technologically central patents in the semiconductor industry decreases in the periods 1991-1995 and 2001-2006, compared to the previous periods. This is due to the decrease in the number of assignees, which in turn is caused by a fall in the number of marginal innovators in those two periods. The number of core innovators is relatively stable over time. More precisely in the first half of the 1990s and the 2000s we observe a peak in of innovative exit and the effect on the total number of innovators is further exacerbated by the fall of innovative entrance. Therefore, in these two periods the industry has undergone a phase of innovative shakeout. As we explained in Chapters 2 and 3, for the first period the cause was a significant shift in the main engineering trajectories caused by a few path-changing inventions that explored new designs for integrated circuits. For the second period, the reason was a change in the ranking of engineering problems induced by the emergence of new technological sub-domains related to LCD and MOSFET technologies.

The data that we have presented so far describes the entire NMPs sample. However, as we discussed in the theoretical framework, in this study we analyse persistence as core innovators. Marginal innovators are not central players in the industry. Therefore, they might exit for reasons that go beyond their knowledge breadth and depth. For this reason, we excluded them from the sample used for the regression analyses. Furthermore, we have to avoid type II errors, i.e. exit misidentification. As in any high-tech industry, semiconductors business are often acquired or merged and companies frequently engage in joint ventures. For our purposes, we need to distinguish real innovative exit from exit by acquisition or merge. In the latter case the acquired or merged organization clearly remains technologically influent if is still found in the list of core innovators. Therefore, if the new entity is considered a core innovator we count it as survival. Else, we consider it as a real exit. There is one of the former cases and ten of the latter in the NMPs sample. Finally, some cases organizations temporarily exit from the list of core innovators

and join it back later. In this case, we still count this as a real exit because, even if momentarily, the given organization lost its technological centrality. We found twelve of such cases in the NMPs sample.

**Table 14: Size of the NMPs sample and categories of innovators**

| Time period   | 1981-1985 | 1986-1990 | 1991-1995 | 1996-2000 | 2001-2006 |
|---|-----------|-----------|-----------|-----------|-----------|
| Number of technologically influent patents ( <i>size of the NMPs sample</i> ) | 689       | 818       | 661       | 1067      | 901       |
| Share of total USPTO semiconductor patents                                    | 42.64%    | 32.45%    | 15.30%    | 17.63%    | 9.36%     |
| Number of assignees ( <i>i.e. innovators</i> )                                | 96        | 145       | 120       | 167       | 132       |
| Core innovators   | 64        | 89        | 87        | 89        | 77        |
| Marginal innovators   | 32        | 56        | 33        | 78        | 55        |
| New innovators ( <i>from previous period</i> )                                | 42        | 88        | 54        | 102       | 74        |
| Incumbent innovators  | 54        | 57        | 66        | 65        | 58        |
| Exit innovators ( <i>from previous period</i> )                               | 48        | 39        | 79        | 55        | 109       |

Moreover, out of the six periods for which we computed the NMPs we can only use four for the empirical analysis. This is because we need to identify new innovators by discriminating prior experience, as the latter is one of the control variables that we include in the logit model described in the next section. The period 1976-1980 is used to provide information on prior experience for the first period used, i.e. 1981-1985. Moreover, our data are right censored. We need to truncate the sample to 2000, in order to be able to observe survival to the next period. This is because we do not have information on which firms survive after 2006. To sum up, the panel dataset that we use for the econometric analysis is made by 133 unique core innovators followed over four five-year periods: 1981-1985, 1986-1990, 1991-1995, 1996-2000. The 133 core innovators hold 2862 patents between 1980 and 2000.

We classified core innovators according to their business type. We mostly relied on data from ICinsights, a business consulting firm specialized in the semiconductor industry.<sup>20</sup> When needed we complement it with additional information gathered from company websites or Wikipedia web-pages. Assignees that represented subsidiaries belonging to the same company have been manually merged into a single entity. Table 15 provides more details on the functional and geographical composition of core innovators. The total number of core innovators per period per business type is highlighted in bold characters. We then break business types by country of origin (we report information only for the set of key countries in the industry: US, Japan, S.Korea, Taiwan and Singapore<sup>21</sup>). Integrated Device Manufacturers (*IDMs*) are large vertically integrated players that perform all the steps of the value chain (from design to commercialization) in-house. To the contrary, *fabless* and *foundries*, specialized in design and manufacturing, respectively, and often interact with each other. *Suppliers* provide *IDMs*, *fabless* and *foundries* with equipment and materials.

<sup>20</sup> We thank Roberto Fontana for sharing the data.

<sup>21</sup> Note that, up to 2006 no single patent from Chinese firms managed to make it in the NMPs sample.

**Table 15: Core innovators' functional and geographic composition**

|  | 1981-1985 | 1986-1990 | 1991-1995 | 1996-2000 | 2001-2006 |
|--|-----------|-----------|-----------|-----------|-----------|
| Equipment suppliers                        | 1         | 2         | 2         | 5         | 9         |
| JP   | 1         | 1         | 1         | 3         | 3         |
| US   |           | 1         | 1         | 2         | 5         |
| Fabless                                    |           | 4         | 4         | 5         | 5         |
| TW   |           |           |           | 1         | 1         |
| US   |           | 4         | 4         | 4         | 4         |
| Government or Industry Research Institutes | 6         | 8         | 9         | 11        | 8         |
| JP   | 1         | 3         | 2         | 2         | 2         |
| KR   |           | 1         | 2         | 2         | 1         |
| TW   |           | 1         | 2         | 2         | 1         |
| US   | 3         | 3         | 2         | 4         | 3         |
| Integrated Device Manufacturers            | 39        | 53        | 50        | 46        | 37        |
| DE   | 1         | 1         | 1         | 2         | 2         |
| JP   | 10        | 14        | 14        | 16        | 12        |
| KR   |           | 1         | 3         | 3         | 4         |
| TW   |           |           | 1         | 3         | 4         |
| US   | 23        | 30        | 25        | 17        | 13        |
| Material suppliers                         | 1         | 2         | 4         | 5         | 5         |
| DE   |           |           |           | 1         | 1         |
| JP   |           | 1         | 3         | 2         | 2         |
| US   | 1         | 1         | 1         | 2         | 2         |
| Pure-Play Foundries                        |           |           | 3         | 4         | 4         |
| SG   |           |           | 1         | 1         | 1         |
| TW   |           |           | 2         | 3         | 3         |
| Universities                               | 3         | 3         | 2         | 4         | 3         |
| US   | 3         | 3         | 2         | 4         | 3         |
| Users                                      | 14        | 17        | 13        | 9         | 6         |
| DE   |           |           | 1         |           | 1         |
| JP   | 1         | 2         | 2         | 2         | 1         |
| US   | 12        | 14        | 9         | 6         | 3         |

*Users* are players external to the industry whose products incorporate semiconductor devices. Some users invest in R&D for specific needs related to semiconductor technologies. Furthermore, there are a number of universities, government and industry research centres that are particularly active in the R&D landscape of the industry, we will lately collectively refer to them in this paper as *research providers*. In terms of number of core innovators, it seems that innovation was demand-pulled in the 1980s and become more and more technology-pushed toward the 2000s. Indeed the importance of downstream innovation (by users) is reducing, in favour of upstream-oriented innovation (by equipment and material suppliers) over time.

### 5.3.2 Identification of Semiconductor technology domains

To measure knowledge breadth and depth of core innovators we first need to identify the set of semiconductor technology domains. Our definition of domains as areas of applied research that share a set of common technological problems which are tackled applying similar approaches, make it possible to exploit the characteristics of the patent system to empirically identify them. Patent claims describe the solution to given engineering problems, and the reference list highlights on which other patents the claimed invention has improved upon. The presence of citations between patents proves that similar mindsets and toolboxes have been applied to tackle the given problems, as discussed in the USPTO Manual of Patent Examining Procedure (USPTO, 2014). Therefore, to empirically identify semiconductor technology domains we can partition the network of patents into communities using a method proposed by Newman (2004) based on modularity maximization. The method identifies areas of the network (called communities) whose nodes (patents in our case) are more related to each other than with nodes belonging to different communities. Patent relatedness is defined by the presence of citations. As such, a community identifies a region of the technological space whose patents are closely related by sequences of technical improvements. This distinguishes them from patents located in other regions of the citation network. We can then interpret communities as technology domains within the meso-family of semiconductor technologies. To identify network communities Newman proposes to maximize a measure of *modularity* ( $Q$ ) defined as follows:

$$Q = \sum_k (e_k - a_k^2) \quad (5.1)$$

Where  $e_{ii}$  is the fraction of edges falling within community  $k$  and  $a_k^2$  is equal to the squared sum of edges falling between communities, as  $a_k = \sum_{j \neq k} e_{kj}$ . As explained by Newman (2004), to assess the significance of the community structure the fraction of edges that fall within communities is compared to the expected value of the same quantity if edges would fall at random without regard for the community structure. If a particular partition returns no more within-community edges than would be expected by random chance modularity  $Q$  would be equal to zero. The best partition is found by iteratively optimizing modularity. The optimization approach starts from the worse possible combination (each node is a community) and then begins an iterative aggregation process which stops when the increase of modularity becomes negative. The modularity maximization procedure and the comparison with equivalent random networks returns the best partition of the network analysed, without assuming a pre-existing community structure. Therefore,  $Q$  is not only a measure of the quality of the partition but can be effectively used to measure the interdependence between domains. Larger values of  $Q$  mean that the different domains of the technological space are less interdependent as the identified partition of the network is more statistically significant (i.e. boundaries across domains are more strongly defined). Note that, to apply the algorithm to the case of directed networks we first need to make the relationships across patent pairs symmetric. This is done by transforming citations into undirected edges. This reinforces the interpretation of citation as a measure of technological relatedness between pairs of patents.

We feed the community detection algorithm with the set of key semiconductor patents granted within each of the following periods 1976-1985, 1976-1990, 1976-1995, 1976-2000, 1976-2006. By cumulating periods, we are able to assess variation on the modularity structure of all generations of semiconductor patents caused by the entrance of young patents. This cumulative perspective takes into account the possibility that in some cases young patents might not



be directly related to each other but still be part of the same domain if they have a strong genetic proximity with the same ancestors. One possible criticism to this approach is that Newman's modularity maximization does not consider the temporal structure of a citation network when evaluating the statistical significance of an existing link between a pair of nodes compared to the probability that the given link exists in a random network. As shown by Jaffe and Trajtenberg (2007), patents have a relatively short forward citation life. Typically, the number of yearly citation to the average patent starts decreasing after 3-4 years from the grant date. The community detection algorithm does not consider this when evaluating the significance of the presence/lack of a citation between a pair of patents compared to a random network in which links among pairs of nodes fall completely at random. Whether this should be considered as a serious bias depends on the extent to which the algorithm tends to identify domains purely based on the age composition of the patents they are composed of. In the Appendix A.4.3, we showed that the age structure of the semiconductor domains identified with the Newman's algorithm is highly scattered. This proves that in our case the citation lag bias is not strong enough to undermine the validity of the domains identified by the community detection process. Table 16 reports basic statistics of the domains identified by Newman's algorithm.

**Table 16: Basic statistics for the identified semiconductor technology domains**

|                                 | 76-85  | 76-90  | 76-95  | 76-00  | 76-06  |
|---------------------------------|--------|--------|--------|--------|--------|
| Number of domains               | 15     | 14     | 29     | 15     | 29     |
| Size of largest domain          | 328    | 368    | 272    | 637    | 701    |
| % of patents in largest domain  | 21,30% | 13,74% | 9,31%  | 13,98% | 11,12% |
| Size of smallest domain         | 29     | 52     | 20     | 62     | 53     |
| % of patents in smallest domain | 1,88%  | 1,94%  | 0,68%  | 1,36%  | 0,84%  |
| Average domain size             | 102,66 | 191,29 | 100,79 | 303,80 | 217,44 |
| Patent dispersion (1-HHI)       | 0,9591 | 0,9873 | 0,9837 | 0,9733 | 0,9868 |

The patent dispersion across domains is measured by the Herfindal-Hirshman Index (HHI), normalized by the maximum possible value of dispersion attainable given the number of domains (similarly to the knowledge breadth index explained in Section 5.2.1.4). Values closer to 1 indicate maximum dispersion across domains, thus pointing to a perfectly equal distribution of patents across communities. The table reveals high patents dispersion. This shows that the community structure is not dominated by a few domains. This further proves the meaningfulness of the identified network partition.

## 5.4 Model

We estimate, using maximum likelihood, the following pooled data logit model, where the probability of survival for observation  $i$  belonging to core innovator  $j$  is:

$$P(s)_{ij} = P(S_{ij} = 1 | X_{ij}, X_i, \alpha) = F(X_{ij} + X_i + \alpha) \quad (5.2)$$

$S_{ij}$  is a binary variable taking value of one if player  $j$  persist being a core innovator in the following period. The matrix  $X_{ij}$  is composed of the list of predictors of innovative survival for each observation  $i$  belonging to player  $j$ . These predictors are: knowledge breadth (KB), knowledge

depth (KD), count of NMPs patents (PC), number of periods of experience as core innovator and business model followed by the innovator. The latter is a factor variable that distinguishes between IDM, foundry, fabless, supplier, research provider and user.  $X_i$  provides information that is player-invariant, in our case it is the level of knowledge persistence in the period in which observation  $i$  has been recorded. Finally,  $\alpha$  serves as a baseline probability of survival, which is included to capture the effect that core innovators have an intrinsic propensity to survive.

Since we are treating time variant data at the firm level as pooled cross-section, we estimate robust error by clustering the variance-covariance matrix by core innovator. Note that core innovators that survived for more than one period entered in the pooled cross-sectional sample twice or more. Clustering the variance-covariance matrix by core innovator allows for arbitrary correlations across observations belonging to the same player, leaving the form of the correlation free to vary from player to player. For the sake of comparison we also estimate a panel data version of the model described in (1) with innovator-specific random effects<sup>22</sup>. Random effects model are more efficient than their cluster pooled data counterpart because the firm-specific effect is estimated separately from the error term. However, random effect models assume that the within-firm correlation takes the same form for all innovators. We recognize that the assumption behind random effects might not be fully legitimate in this context. Firms pursue different strategies that interact with the level of their capabilities. This might take different forms and, therefore, differently affect the probability of survival. Furthermore, it might also create correlation between the individual random effect component of the error term and the KB and KD regressors. Nevertheless, the rationale of comparing the results with a random effect model here has the specific purpose of investigating whether the overall results holds and whether we are able to reduce the variance of the predicted probability of survival for fully specialized and poorly diversified players. In the panel data random effect version of model (1) the probability of survival for observation  $i$  belonging to core innovator  $j$  at time  $t$  is:

$$P(s)_{ijt} = P(S_{jt} = 1 | X_{ijt}, X_j, X_t, \mu_j) = F(X_{ijt} + X_j + X_t + \mu_j) \quad (5.3)$$

Where  $X_{ijt}$  are time dependent predictors for observation  $i$  belonging to player  $j$  (i.e. our KB, KD, PC and Experience variables),  $X_{ij}$  is a time independent predictor that varies across players  $j$  (i.e. the business model variable),  $X_t$  is a time dependent predictor that is constant across players (i.e. knowledge persistence) and  $\mu_j$  is the core innovator specific effect which represents the combined effect of all omitted subject-specific unobserved variables that cause some subjects to be more (or less) more likely to survive than others.

To test the robustness of our findings to the estimation method used we also run probit and complementary log-log versions of the pooled and panel data models described in (1) and (3). All the main finding hold regardless of the method used to linearize the relationship between survival and the predictors. Therefore, we will only discuss findings for the logit model and report results for the probit and complementary log-log versions in the appendices A.5.4 and A.5.5.

---

<sup>22</sup> We opted for a random effect model instead of a fixed effect one as the latter severely reduces the sample size. The estimation of a fixed effect logit model relies on conditional maximum likelihood, therefore it necessarily drops 80 right censored observations (i.e. those core innovators that never exit over the given window of time) because the dependent variable is a constant in this case. This reduces the sample size by 207 observations. The effect of such a large harvest is very problematic as the maximum likelihood maximization process does not converge for most of our models and the estimation becomes impossible.

Given the stability of findings, we opted to present the findings for the logit model instead of the probit or complementary log-log versions because of the ease to interpret coefficients as odds ratios provided by the former. Note that, since we include a variable measuring innovators' periods of experience and a baseline survival probability  $\alpha$ , the complementary log-log version of our model becomes conceptually similar to the discrete time representation of an underlying continuous time proportional hazard model. This is particularly convenient as it means that our results would be confirmed by a formally defined discrete-time survival analysis.

## 5.5 Measurement of the key variables

To study the effects of firms' technical knowledge diversification and accumulation and the stability of the engineering trajectory on firms' innovative survival, we need to define measures of firms' knowledge breadth, depth and persistence of technical knowledge. We also need a measure of knowledge modularity to assess the difficulty of knowledge diversification. In the literature knowledge diversification has been measured by counting active fields of research, by computing the concentration of a firm's patents across technological classes or by calculating firms' specialization indices across fields (e.g. Xu, 2014; D'Este, 2005; Brusoni et al., 2005; Brusoni and Geuna, 2003). Knowledge accumulation is usually measured by firms' patent counts in each class (e.g. Xu, 2014) or by counting areas of fields in which firms are co-specialized in basic and applied research. (e.g. Brusoni et al., 2005). These approaches dissatisfy us for a number of reasons. First, as we discussed in Section 5.3.2, relying on classification systems to define technology domains does not consider that technology relatedness goes beyond class boundaries. Furthermore, classifications react with a long lag to changes in the structure of the system of technologies<sup>23</sup>. Bottom-up techniques, such as community detection on citation networks, identify domains in *real time*, without assuming any prior knowledge of relatedness across domains. In this sense, they are not affected by any cognitive or time bias. This is particularly important to detect the level of knowledge diversification of surviving firms. Moreover, measuring diversification by looking at the distribution of item counts across domains does not provide any information on how much of the domain-specific knowledge a firm actually mastered. Similarly, having the same amount of patents in two or more classes does not necessarily imply that the firm's technical knowledge is equally distributed across those classes. This is because the difficulty of generating patentable inventions might vary greatly across domains. Nor does it provide any information on how the inventions disclosed in the firm's patents are related to the existing prior art. Because of these reasons, we need measures of technical knowledge diversification, accumulation and persistence that respect the structural properties of the knowledge system and its topography. To define a suitable measure for each of these dimensions of technical knowledge we rely on the similarity between knowledge and genetic evolution. Technological knowledge largely builds on past achievements and firms innovative efforts draws on prior internally and externally developed technological solutions. As such, technological knowledge is

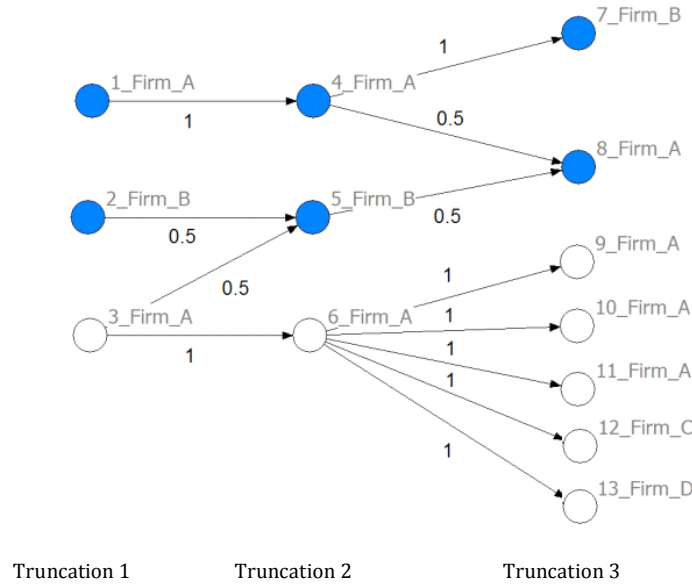
---

<sup>23</sup> For instance, the USPTO introduced a cross-reference art collection for Nanotechnology in October 2004. According to the output of the USPTO reclassification efforts of old patents into the newly created art collection class, the first Nanotechnology patent has been granted in 1978. This means that it took twenty-six years for a social agreement on the existence of the Nanotechnology domain to emerge and lead to the creation of an ad-hoc technology class. This is because the patent system is a complex engineering system rich in technology and human complexity (de Weck et al., 2011).

largely cumulative, although from time to time radical innovations are introduced. These cause a shift in the technological trajectory and, if successful, permanently change the direction of technological change. When this happens, in order to survive firms need to adapt their capabilities to the new technological environments (unless, of course, if they are those that introduced the radical innovations). An appropriate analogy with genetic evolution can be easily made. Populations also evolve slowly and in a cumulative fashion, as genetic traits are inherited from generation to generation. Shocks in the form of random mutations can break this persistency by introducing novel genetic traits, whose survival will ultimately be decided by natural selection. A second source of genetic variety takes the form of transmission from parents to offspring of genes that evolved as the result of adaptation to the environment (i.e. Lamarckian inheritance). This analogy inspired the definition of a genetic approach (GA) to patent citation networks, which has been developed by Martinelli (2010) and Martinelli and Nomaler (2014). The approach provides an original way to measure technological knowledge, its persistence and evolution. Our main contribution to the genetic approach lays in the definition of a method to map firm technological genome and its genetic heritage. In addition, we also define a set of metrics that allow applying this approach to measure firm's technical knowledge breadth, depth and knowledge persistence at the industry level. The GA looks at patents as knowledge genes. Citations therefore represent channels through which knowledge is *inherited* from cited to citing patents. Therefore, a citation network can be interpreted as a map of genetic history, portraying knowledge that persists from parents (cited patents) to offspring (citing patents). In the following, we describe how we construct our measures of firms' knowledge breadth, depth, and knowledge persistence and modularity at the patent level.

### 5.5.1 Knowledge persistence

Figure 3 represents a dummy citation network. Nodes stand for patents and arcs represent citations from the patent to the right to the one on the left (the arrow indicates the direction of improvement). The first step of the patent genetic decomposition requires sorting the network topologically, meaning that nodes are ordered such that for every directed citation  $ab$  from patent  $b$  to patent  $a$ ,  $a$  comes before  $b$  in the ordering. Thus, network truncations can be identified. Truncations are analogous to generations in a genealogy family tree. In Figure 32 we have three truncations. The number of truncations corresponds to the largest possible path in the network. Nodes labels report a fictitious patent number and the name of the assignee. Nodes' colour highlights fictitious technological domains.



**Figure 32: Fictitious citation network with shares of genetic heritage**

Let us assume that filled nodes represent patents belonging to domain A and white ones belong to domain B. The weight of the arcs represents the share of knowledge *direct genetic heritage* (we refer to it as the weight  $w$ ), meaning the share of the knowledge included in the citing patent that is *inherited* from the cited one. The weight between patents  $i$  and  $j$  is computed as follows:

$$w_{ij} = \frac{cit_{ij}}{bw d cit_j} \quad (5.4)$$

Where  $cit_{ij}$  is a dummy variable taking the value of one if  $j$  cites  $i$  and zero otherwise and  $bw d cit_j$  is the number of backward citations made by patent  $j$ . For instance, patent 7 cites patent 4 only. Therefore 100% of the knowledge that patent 7 inherits from the set of patents belonging to the directly preceding truncation comes from patent 4. Patent 8, instead, builds on patents 4 and 5, meaning that these two patents contribute 50% each to the first layer of inherited knowledge of patent 8. However, the genetic decomposition of the citation network goes beyond direct citations and keeps into account the topology of the whole network (i.e. all the connection across all 'generations'). Indeed, as explained by Martinelli and Nomaler (2013), the genetic approach decomposes the knowledge content of a given node in function of the nodes that precede it. Therefore, we need to account for genetic heritage across all possible truncations. Following the same example, we can see that 100% of the knowledge inherited by patent 4 from the preceding truncation comes from patent 1, whereas 50% of the knowledge of patent 5 comes from patent 2 and 50% from patent 3. Consequently the contribution of patent 1 to patent 8's knowledge is equal to 0.5 ( $=1*0.5$ ), whereas patents 2 and 3 contributes 0.25 ( $=0.5*0.5$ ) to patent 8's knowledge. It follows that we can define an index of the *total genetic heritage* (from now on just *heritage*) from each patent to each other one. The *heritage* index between a pair of patents  $i$  and  $j$  is computed multiplying the weights of all the links connecting patents  $i$  and  $j$ . The *heritage* index can be easily normalized such that the sum of the weights of the backward citations originated from a given patent adds up to one. In this case, the index becomes the *share of the total genetic heritage* (*normheritage*) of the citing patent that is inherited from each of the other patents in the network.

For the sake of clarity we report in Table 17 the *heritage* and the *normheritage* (in italics and as percentage) for each pair of patents in the dummy network of Figure 32. It is important to notice that, as argued by Martinelli and Nomaler (2014), the genetic decomposition does not underplay the role of novelty. One has to distinguish between knowledge that is inherited and knowledge that is generated. Citation weights represent shares of the former. The existence of the citation represents the latter, as citations identify prior art which is subsequently improved. Therefore, the very same existence of the citation proves the relevance of the novelty created by the cited patent.

We can now calculate the *persistence index* (PI) for patents belonging to the first two truncations in the network of Figure 3. For this purpose let us define the set  $E$  of endpoint patens  $i$  such that:

$$E = \{ i \mid bwdcit_i > 0, fwdcit_i = 0 \} \quad (5.5)$$

Where  $bwdcit_i$  and  $fwdcit_i$  are backward and forward citation made and received by endpoint patent  $i$  respectively. Endpoints are therefore those patents that end the sequence of citations in the network. We index them by  $e \forall i \in E$ . The *persistence index* (PI) computes how much of the knowledge of a given patent  $i \notin E$  is retained in the set of endpoints  $E$ . It can be computed as follows:

$$PI_i = \sum_e heritage_{ie} \quad (5.6)$$

The PI for the patents of the fictitious network of Figure 32 are reported in the last raw of Table 18.

**Table 17: Measuring knowledge genetic heritage across patents**

|   |                          | Parents patent ( <i>i.e. cited</i> ) |                      |                       |                      |                      |                      |                |
|---|--------------------------|--------------------------------------|----------------------|-----------------------|----------------------|----------------------|----------------------|----------------|
|   |                          | 1                                    | 2                    | 3                     | 4                    | 5                    | 6                    | Total heritage |
| Offspring patents<br>( <i>i.e. citing</i> ) | 4                        | 1.00<br><i>100.0%</i>                | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>   | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 1              |
|   | 5                        | 0.00<br><i>0.0%</i>                  | 0.50<br><i>50.0%</i> | 0.50<br><i>50.0%</i>  | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 1              |
|   | 6                        | 0.00<br><i>0.0%</i>                  | 0.00<br><i>0.0%</i>  | 1.00<br><i>100.0%</i> | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 1              |
|   | 7                        | 1.00<br><i>50.0%</i>                 | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>   | 1.00<br><i>50.0%</i> | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 2              |
|   | 8                        | 0.50<br><i>25.0%</i>                 | 0.25<br><i>12.5%</i> | 0.25<br><i>12.5%</i>  | 0.50<br><i>25.0%</i> | 0.50<br><i>25.0%</i> | 0.00<br><i>0.0%</i>  | 2              |
|   | 9                        | 0.00<br><i>0.0%</i>                  | 0.00<br><i>0.0%</i>  | 1.00<br><i>50.0%</i>  | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 1.00<br><i>50.0%</i> | 2              |
|   | 10                       | 0.00<br><i>0.0%</i>                  | 0.00<br><i>0.0%</i>  | 1.00<br><i>50.0%</i>  | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 1.00<br><i>50.0%</i> | 2              |
|   | 11                       | 0.00<br><i>0.0%</i>                  | 0.00<br><i>0.0%</i>  | 1.00<br><i>50.0%</i>  | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 1.00<br><i>50.0%</i> | 2              |
|   | 12                       | 0.00<br><i>0.0%</i>                  | 0.00<br><i>0.0%</i>  | 1.00<br><i>50.0%</i>  | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 1.00<br><i>50.0%</i> | 2              |
|   | 13                       | 0.00<br><i>0.0%</i>                  | 0.00<br><i>0.0%</i>  | 1.00<br><i>50.0%</i>  | 0.00<br><i>0.0%</i>  | 0.00<br><i>0.0%</i>  | 1.00<br><i>50.0%</i> | 2              |
|   | <i>Persistence Index</i> | 2.50<br><i>14.7%</i>                 | 0.75<br><i>4.4%</i>  | 6.75<br><i>39.7%</i>  | 1.50<br><i>8.8%</i>  | 0.50<br><i>2.9%</i>  | 5.00<br><i>29.4%</i> | 17             |

In our example, the most prominent patent in terms of persistence is patent 3, which accounts for 37.5% of the knowledge retained in the last truncation of patents. The example we used so far represents a static picture of a citation network. By dynamically analysing different snapshots of subsequent networks, it is possible to measure how knowledge persistence changes over time. Suppose that we observe the citation network for three periods,  $T - T+C$ ,  $T+C - T+2C$  and  $T+2C - T+3C$ . We can measure *recent knowledge persistence* from one period to the next by computing the persistence index including only patents granted within one period, take for instance period  $T+C - T+2C$ , in the list of parents patents and only patents granted in the next period,  $T+2C - T+3C$ , in the list of endpoints. The ratio of the total heritage of endpoints belonging to the third period received from parent patents granted in the second is a measure of recent knowledge persistence from period 2 to 3. Similarly, we can measure *cumulative knowledge persistence*, for the third period, by including in the list of parent nodes all patents granted in the first and second period. Obviously, this version of the index tends to decrease over time as generation of patents goes by.

**Table 18: Measuring knowledge persistence**

|                   |    | Parents patent (i.e. cited) |       |       |       |       |       |                |
|-------------------|----|-----------------------------|-------|-------|-------|-------|-------|----------------|
|                   |    | 1                           | 2     | 3     | 4     | 5     | 6     | Total heritage |
| Endpoints         | 7  | 1.00                        | 0.00  | 0.00  | 1.00  | 0.00  | 0.00  | 2              |
|                   |    | 50.0%                       | 0.0%  | 0.0%  | 50.0% | 0.0%  | 0.0%  |                |
|                   | 8  | 0.50                        | 0.25  | 0.25  | 0.50  | 0.50  | 0.00  | 2              |
|                   |    | 25.0%                       | 12.5% | 12.5% | 25.0% | 25.0% | 0.0%  |                |
|                   | 9  | 0.00                        | 0.00  | 1.00  | 0.00  | 0.00  | 1.00  | 2              |
|                   |    | 0.0%                        | 0.0%  | 50.0% | 0.0%  | 0.0%  | 50.0% |                |
|                   | 10 | 0.00                        | 0.00  | 1.00  | 0.00  | 0.00  | 1.00  | 2              |
|                   |    | 0.0%                        | 0.0%  | 50.0% | 0.0%  | 0.0%  | 50.0% |                |
|                   | 11 | 0.00                        | 0.00  | 1.00  | 0.00  | 0.00  | 1.00  | 2              |
|                   |    | 0.0%                        | 0.0%  | 50.0% | 0.0%  | 0.0%  | 50.0% |                |
|                   | 12 | 0.00                        | 0.00  | 1.00  | 0.00  | 0.00  | 1.00  | 2              |
|                   |    | 0.0%                        | 0.0%  | 50.0% | 0.0%  | 0.0%  | 50.0% |                |
|                   | 13 | 0.00                        | 0.00  | 1.00  | 0.00  | 0.00  | 1.00  | 2              |
|                   |    | 0.0%                        | 0.0%  | 50.0% | 0.0%  | 0.0%  | 50.0% |                |
| Persistence Index |    | 1.50                        | 0.25  | 5.25  | 1.50  | 0.50  | 5.00  | 14             |
|                   |    | 10.7%                       | 1.8%  | 37.5% | 10.7% | 3.6%  | 35.7% |                |

### 5.5.2 Firms' knowledge breadth and depth

To develop a measure of firms' knowledge breadth and depth we first need to define an index of knowledge proximity that can be used to identify the position of a given firm in the multi-domain technological space (i.e. the genetic composition of its technological knowledge base). Let us define the set of patents belonging to firm  $f$  as  $F$  and the set of patents belonging to technology  $k$  as  $K$ . We then define the set  $Z$  as the intersection of  $F$  and  $K$ :

$$Z = F \cap K = \{i \mid i \in F \text{ and } i \in K\} \quad (5.7)$$

The *proximity* index between a given firm  $f$  and the set of technological domains  $k$  is then calculated as the total genetic heritage of firm  $f$  coming from domain  $k$ , as shown in the following equation.

$$proximity_{fk} = \sum_{i,j \in Z} heritage_{ij} \quad (5.8)$$

Table 19 reports the firm-technology domain proximity for our fictitious network. Firm A has the largest knowledge base, as measured by its total heritage. Twenty percent of its technical knowledge is related to technology domain 1 and eighty percent to domain 2.

**Table 19: Measuring firm-technology domain proximity**

|        | Cited  |               |                |
|--------|--------|---------------|----------------|
|        | Tech 1 | Tech 2        | Total heritage |
|        |        |               |                |
| Citing | Firm A | 1.75<br>20.0% | 7<br>80.0%     |
|        | Firm B | 3.5<br>82.4%  | 0.75<br>17.6%  |
|        | Firm C | 2<br>0%       | 2<br>100%      |
|        | Firm D | 2<br>0%       | 2<br>100%      |

The *knowledge breadth* (KB) index for a firm  $f$  is calculated as the Herfindal-based dispersion index of a firm's proximity to the different technological domains  $k$  normalized by the maximum Herfindal index attainable given the total number of technological domains (equal to  $K$ ). The equation is as follows.

$$KB_f = \frac{1 - \sum_{k=1}^K proximity_{fk}^2}{1 - \left(\frac{1}{K}\right)^2 \cdot K} \quad (5.9)$$

Where the denominator is equal to the highest possible value the numerator can attain. *Proximity* is calculated using *normheritage*. The KB index ranges from a zero to one. A value of zero indicates that firm  $f$  is fully specialized, meaning that 100% of its knowledge is retained from a single technological domain. To the contrary, a value of one means that firm  $f$  is fully diversified, i.e. its knowledge is equally spread across all existing domains.

The *knowledge depth* index (KD) measures how much of the knowledge persistence of the given domains in which firm  $f$  operates is retained by the firm. In other words, KD measures how knowledgeable the firm is in the domains in which it operates. The index is calculated as follows.

$$KD_f = \frac{\sum_{k=1}^{K_f} \left( proximity_{fk} / \sum_{f=1}^F proximity_{fk} \right)^2}{K_f} \quad (5.10)$$

Where *proximity* is calculated using *normheritage* and  $F$  is equal to the number of firms and  $K_f$  is the number of domains in which  $proximity_{fk} > 0$ . The denominator is equal to the largest numerator attainable by firm  $f$  given the number of areas in which it operates. KD ranges from zero to



one. It takes the value of zero when firm  $f$  is a marginal player in each domain in which it operates. Values closer to one indicate that the firm monopolized the domains in which it operates, meaning that 100% of the persistent knowledge in those domains is retained by the firm. It is easy to verify that, in our example, the values of knowledge breadth and depth for our fictitious firms are as follows:  $KB_A=0.64$ ;  $KD_A=0.23$ ;  $KB_B=0.58$ ;  $KD_B=0.22$ ;  $KB_C=0.0$ ;  $KD_C=0.01$ ;  $KB_D=0.0$ ;  $KD_D=0.01$ .

In summary, the *knowledge breadth* index measures the width of a given firm's knowledge tree, i.e. how many roots the patent citation network of the given firm has, whereas the *knowledge depth* measures how long the roots are. Note that, from a strategic point of view, it exists a trade-off between increasing knowledge breadth and depth. Suppose that a successful specialized firm (i.e. one with high KD but narrow KB) decides to diversify into an existing technological domain. Most likely, the firm will face some initial difficulties in mastering the new technology. Consequently, at first it will be less knowledgeable than rival firms that were already present in that domain. Therefore, the firm's attempt to increase its KB is likely to decrease its knowledge depth initially. Depending on the firm's capabilities, it might ultimately be able to diversify successfully or it might need to reconsider its strategy and move back to a specialization position.

Based on the level of KB and KD we can classify core innovators in four categories, as shown in the taxonomy illustrated in Table 20. We classify core innovators according to the following four categories: *Diversified Leaders (DL)*, *Diversified Followers (DF)*, *Specialized Leaders (SL)* and *Specialized Followers (SF)*. We use this classification to test the third hypothesis, discussed in Section 5.2. In order to compute the survival rate for the four categories we need to decide upon a threshold to use to distinguish high and low levels of knowledge depth and broad and narrow knowledge depth. The thresholds used are based on the distribution of knowledge breadth and depth. We discuss them in Section 5.7.

**Table 20: Knowledge taxonomy of core innovators**

|                        |        | Knowledge Depth (KD)        |                           |
|------------------------|--------|-----------------------------|---------------------------|
|                        |        | Low                         | High                      |
| Knowledge Breadth (KB) | Broad  | <b>Diversified follower</b> | <b>Diversified leader</b> |
|                        | Narrow | <b>Specialized follower</b> | <b>Specialized leader</b> |

### 5.5.3 Measures of knowledge modularity

In order to assess whether semiconductor technologies became more or less modular over time we make use of two indicators: *Newman's Q modularity* and *genetic modularity*. The latter is our original contribution. Newman's Q only accounts for the presence of inter-domain linkages but does not measure the strength of information flowing on them. Genetic modularity reveals the level of knowledge relatedness across domains. Using a genealogy analogy, modularity Q sets family borders within a population, genetic modularity measures genetic closeness across families.

We define a measure of modularity that complements Newman's Q and provides information about the genetic relationships between different domains. Newman's Q only takes into account direct citations (i.e. direct relatedness). However even though a pair of domains might

be connected by few citations, in principle, they could still be genetically interconnected if a lot of knowledge flows through those citations. We therefore construct an index of *genetic modularity*. We first need to define a measure of external knowledge usage at the domain level (DEK). This is computed as follows.

$$DEK_j = 1 - \frac{proximity_{jj}}{\sum_{k=1}^K proximity_{jk}} \quad (5.11)$$

The domain external knowledge index (DEK) measures how much of the knowledge of domain  $j$  is inherited from all domains  $k$  other than  $j$ . Genetic modularity is then defined as follows.

$$Genetic\ modularity = 1 - \left( \frac{1}{K} \sum_k DEK_k \right) \quad (5.12)$$

Genetic modularity measures how much of the knowledge in the system comes from inter-domains genetic relationships on average. Genetic modularity ranges from zero to one. Higher values indicate that technology domains are more genetically independent from each other. Values closer to zero reveal knowledge admixture across technology domains.

#### 5.5.4 Control variables

In our model, we include two control variables for the technological size of the firm and its level of innovative experience. We measure the size of core innovators' technical knowledge base as the count of technologically influent patents they have been granted. This simply equals the number of NMPs patents granted to each core innovator in each of the five periods. In the remaining of the paper, we refer to this variable with the acronym PC (patent count). Finally, the number of periods in which each core innovator has showed up in the NMPs sample is used as a proxy of the players' innovative experience.

## 5.6 Trends of knowledge modularity and persistence

The goal of this section is to briefly describe the dynamics of technological change followed by the semiconductor industry over the period between the beginnings of the 1980s until mid-2000s. We describe the trends followed by the industry according to our indices of knowledge modularity and persistence.

Table 21 reports the evolution of modularity  $Q$  and genetic modularity of the NMPs sample as well as the set of measures of knowledge persistence. Values of the genetic modularity are even higher. This shows that not only the different domains are connected by a few citations but also that very little knowledge flows pass through those citations. Trends for both indicators do not reveal large variations over time. Essentially semiconductor technology domains are genetically independent from each other. This reinforces the importance of testing our hypothesis concerning knowledge diversification. Since technical knowledge is not easily portable across domains in the semiconductor industry, diversification needs to be already in place before the occurrence of possible shocks in knowledge persistence as reactive behaviour is hindered by the lack of bridges across domains. The last four rows of Table 21 show the trend in knowledge per-

sistence from 1980 to 2000. Since persistence is calculated from one period to the next, obviously, the value for the last period is not available. We report two versions of the recent and cumulative knowledge persistence indices. The first one is computed for the whole NMPs sample. As we have discussed in Section 5.3, the latter includes the main component of the NMPs for each of the five periods under observation and the second largest component for periods 1991-1995 and 2001-2006. This is done to take into account that a larger than usual share of influential patents were found in the second component in these two periods. We also compute recent and cumulated knowledge persistence on a sub-set of the NMPs sample that only includes the main component. We refer to the latter as the indices of recent and cumulated knowledge persistence on the main trajectories.

When measured using the NMPs sample, both recent and cumulated knowledge persistence indicators show a cyclical trend for knowledge persistence, with peaks in 1990 and 2000 and lower values in 1985 and 1995. If we look at cumulated knowledge, we see that only 21.9% of the knowledge embodied in the endpoints of the network in the period 1986-1990 was inherited by patents granted before 1986. This means that 78% of the knowledge generated in this period is genuinely novel. Therefore, a lot of new knowledge needs to be learned to keep updated. This figure increased to 53.4 in period 1991-1995 (i.e. less than half of the knowledge is new) and then drop to 13.1% in period 1996-2000. Finally, it increases again to 25.1% in 2001-2006 but without reaching the level of 1991-1995. On average, over the four periods, knowledge persistence is 28.35%. Therefore, more than 70% of the knowledge generated in each period is genuinely new, on average. We argue that this is a strong indicator of knowledge-replacing technological change, which creates a strong evolutionary pressure. The trends differ when we measure persistence on the main trajectories only. Recent knowledge persistence becomes monotonically decreasing, whereas cumulated knowledge decreases up to 1995 and slightly increases from period 1996-2000 to 2001-2006. These differences show that the knowledge replacing nature of technical change has been considerably stronger for those domains that are found on the main trajectories of improvements, i.e. those related to the largest areas of research in the industry.

**Table 21: Trends of knowledge modularity and persistence**

|  | 1981-1985 | 1986-1990 | 1991-1995 | 1996-2000 | 2001-2006 |
|--|-----------|-----------|-----------|-----------|-----------|
| Modularity Q                                 | 0.879     | 0.901     | 0.907     | 0.916     | 0.902     |
| Genetic Modularity                           | 0.975     | 0.981     | 0.947     | 0.977     | 0.966     |
| Cumulated Knowledge Persistence (NMPs)       | 0.220     | 0.534     | 0.131     | 0.251     | n.a.      |
| Recent Knowledge Persistence (NMPs)          | 0.150     | 0.338     | 0.091     | 0.153     | n.a.      |
| Cumulated Knowledge Persistence (Main traj.) | 0.220     | 0.184     | 0.131     | 0.144     | n.a.      |
| Recent Knowledge Persistence (Main traj.)    | 0.150     | 0.110     | 0.091     | 0.084     | n.a.      |

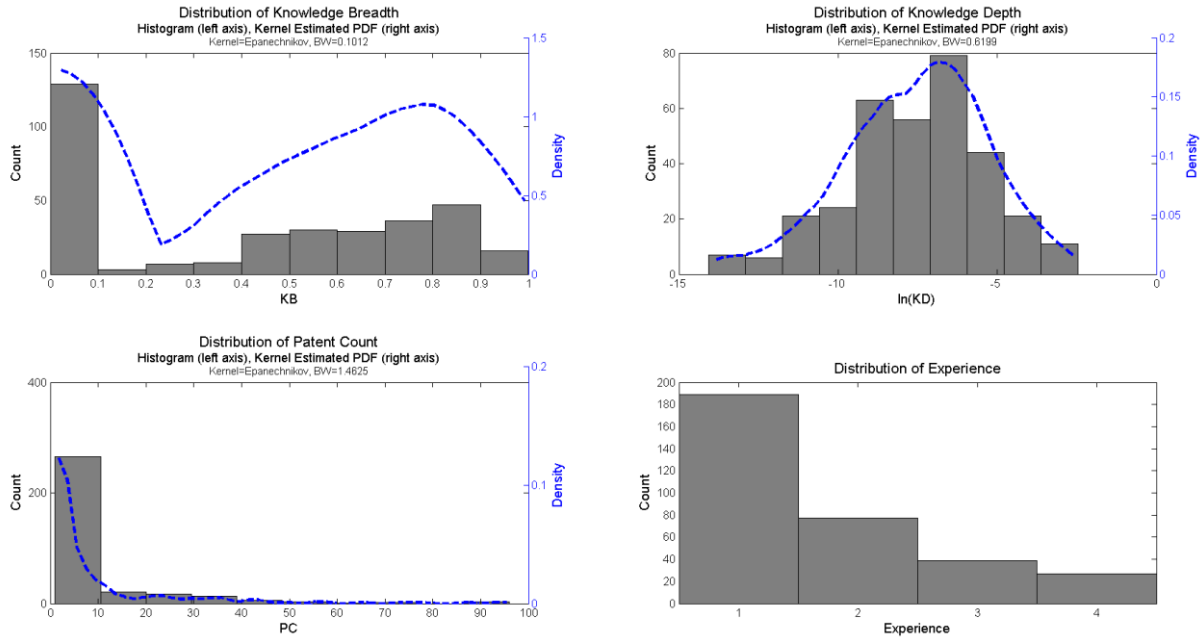
## 5.7 Econometric issues

Before presenting the findings from the logit model estimation, we discuss a few econometric issues that affect the regressions. We also inspect our predictors of innovative survival for possible collinearity.

Figure 33 shows the distribution of four out of the five firm level variables (information about the fifth one, business type, can be found in Table 15, Section 3). Vertical bars represent the number of observations included in each bin, whereas dotted lines show the Kernel smoothed probability density function, estimated with the Epanechnikov method. The distribu-

tion of knowledge breadth, reported in the upper left panel, is clearly bimodal with two peaks, one at 0 and one at 0.77. This indicates that there are two distinct populations of core innovators in terms of knowledge diversification. One is made by fully specialized players and the other is composed of diversified innovators. However, even considering diversified players alone does not return a normal distribution as the right tail is fatter than a normal one. The existence of two distinct populations of core innovators with respect to their level of diversification can potentially make it difficult to assess how knowledge breadth affects the probability of innovative survival. In particular, the large density of observation in the left tail of the distribution followed by a particularly low density for intermediate levels of KB in the range  $0 < KB \leq 0.4$ , is a source of concern for the efficiency and reliability of the estimated effect of KB. This makes the use of a continuous variable of KB particularly unreliable because the estimation of the effects of KB on survival is very inefficient for that range (i.e. the standard deviation of the parameter is very high). For this reason, we build a categorical variable based on ten equally spaced bins of KB. This approach assumes that the effect of the relationship between the independent and the dependent variable is flat within intervals. However, using ten bins greatly mitigate the influence of this assumption. Furthermore, it allows avoiding any prior assumption on the functional form of the relationship by letting the data speak for it. It does not completely solve the problem of unequal density of observations but allow ignoring it by refraining to make any conclusion regarding the effect within the range  $0 < KB \leq 0.4$  and by erasing its influence on the estimation of the effect for the other bins. As a comparison with a continuous version of KB, we also test the effect of diversification by using a restricted cubic spline function. This approach allows maintaining a smooth relationship between KB and survival and maximizes power (by using fewer degrees of freedom than the categorical variable approaches). The spline function transforms the predictor (KB) to achieve linearity while still estimating a non-linear relationship in a smooth way.

The measure of knowledge depth does not present any particular econometric issue. The distribution of KD is log-normal, with a geometric mean of  $0.00052 = e(-7.56)$ . This show how difficult is to cumulate knowledge. On average, core innovators in the semiconductor industry know 0.052% of the total technical knowledge produced in the sub-domain(s) in which they are active. The largest value of KD in our sample is 0.085. This has been achieved by Motorola in the period 1991-1995, when its KB scored 0.95. Table 39, reported in the appendix A.5.1, lists the top 10% observations by KD. Among them, fifty-percent are US innovators, 26.47% are Japanese, 11.76% Taiwanese, 8.82% are Korean and 2.94 are Singaporeans. About 65% are IDMs and circa 15% are foundries. The lower panels of Figure 33, reports the distribution of patent counts (i.e. the number of key patents granted to a given company in a single period) and of periods of experience as core innovator. Both are skewed to the left, with most observations having less than 10 patents and showing up in the NMPs for the first time.



**Figure 33: Distribution of knowledge breadth, depth, patent count and experience**

As discussed in Section 5.5, to test our third hypothesis we construct the categorical variable KNOWCLASS capturing the interaction between KB and KD. Given the nature of the distribution of these two variables, we created two versions of KNOWCLASS. The first one uses the median of KB and of the natural logarithm of KD as cut-off values. As this generates the four categories presented in Section 5.5 we label this version KNOWCLASS4. It follows that for an observation to be considered to have large knowledge breadth and depth it needs to be in the top-half of the distribution for both variables. We also construct a second version of KNOWCLASS (called KNOWCLASS6) which classifies the location of a core innovator in the KB-KD knowledge space into six categories. The boundaries of each category reflect the characteristics of the distributions of KB and KD. In particular we split KB in three groups, low, intermediate and high, using the peaks identified by the kernel estimated probability density function as cut-off values. KD is split using the geometric mean as cut-off value. As such, we can therefore distinguish core innovators according the following six categories, whose boundaries are defined in Table 22: fully specialized followers (FSFs), fully specialized experts (FSEs), mildly diversified followers (MDFs), mildly diversified experts (MDEs), highly diversified followers (HDFs) and highly diversified experts (HDEs). KNOWCLASS6 is the preferred classification as its cut-offs values respect the characteristics of the sample population in terms of KB and KD variable distributions. We need to note that only three observations fall within the HDF category. This is an interesting insight as it tells that highly diversified players also tend to have relatively large levels of knowledge accumulation. However, it also makes the econometric estimation of the probability of survival for HDFs highly unreliable. Therefore, we shall not consider the estimate for this category. The scarcity of observations for core innovators with very large values of KB and very low values of KD, but there are several players with very large KD and low KB. This possibly suggests the presence of unidirectional reinforcing effects of diversification on knowledge accumulation that do not hold in the opposite direction. A possible explanation might be due to a size effect. As firms' technical knowledge base grows, diversification opportunities open up (due to smaller opportunity costs of exploration) and the increased experience allows for knowledge accumula-

tion in several domains. The econometric regressions shall shed light on the presence of size effects.

The last pre-estimation diagnostic that we perform is the correlation and collinearity analysis. Scatter plots of the relationship between continuous explanatory variables (i.e. KB, KD and PC) are shown in Figure 34 and their correlations are reported in Table 23. Although in all three cases the distribution of observations across the variable space follows a general pattern the corresponding correlation is not large enough to create worries in terms of collinearity. In each case, the variable on the horizontal axis explains around 30% of the variance of the variable on the vertical axis. Indeed the variance inflation factors (VIF), reported in Table 23 are low enough to exclude collinearity-induced estimation problems. Panels A and B of Figure 34 are also useful for an early diagnostic of a possible size-effect. We note that all players with a large technical knowledge base are diversified and knowledgeable. However, the variance of KB and KD even at very large levels of PC is sufficiently large to allow distinguishing the effect on survival chances caused by technological size from that caused by diversification or knowledge accumulation in the econometric analysis. This does not hold for the categorical variables KNOWCLASS4 and KNOWCLASS6, as shown by Figure 35. In particular, for values of PC greater or equal than 14 all players have the same level of KNOWCLASS4, whereas the threshold is  $PC \leq 24$  for KNOWCLASS6. This means that regressions will not be able to properly distinguish the effect of size from the effect of diversification when we use these variables. Therefore, we will limit the sample to players with a value of PC lower than the identified thresholds in our regression models.

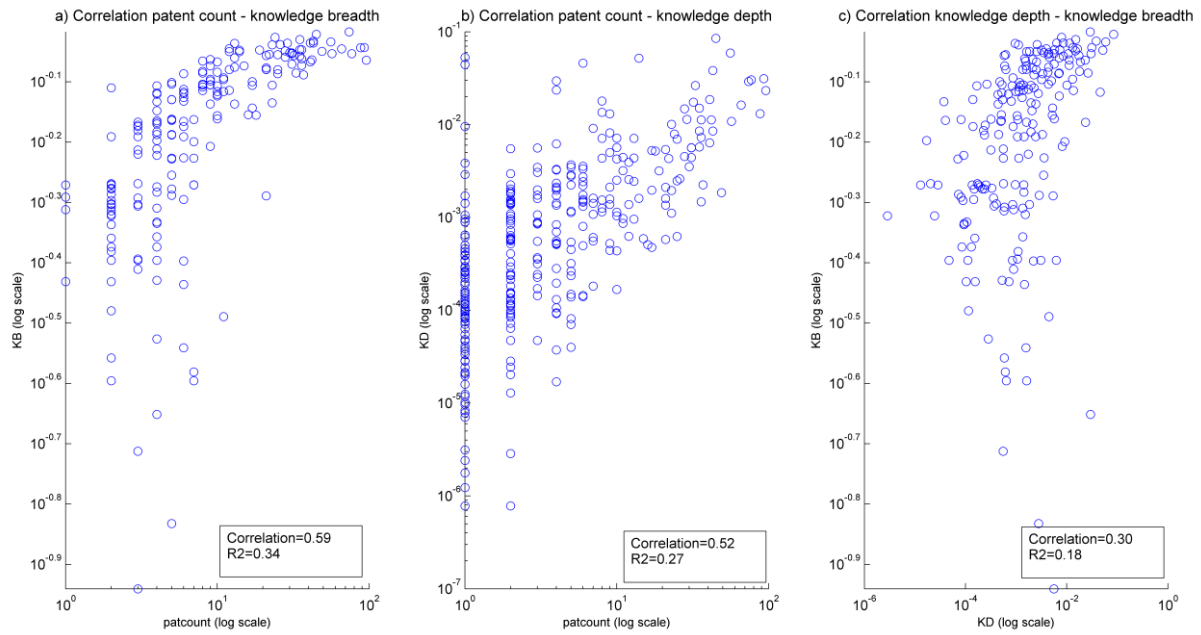
**Table 22: cutoff values of KB cat, KNOWCLASS4 and KNOWCLASS6**

| Variable  | Range   | Count | Freq. |
|---|---|-------|-------|
| <b>Knowledge classification (4)</b><br>(KNOWCLASS4) | Specialized follower (SF): $KB_i \leq M_{KB}$ & $KD_i \leq M_{KD}$                    | 113   | 0.340 |
|   | Specialized expert (SE): $KB_i \leq M_{KB}$ & $KD_i > M_{KD}$                         | 52    | 0.157 |
|   | Diversified follower (DF): $KB_i > M_{KB}$ & $KD_i \leq M_{KD}$                       | 53    | 0.160 |
|   | Diversified expert (DE): $KB_i > M_{KB}$ & $KD_i > M_{KD}$                            | 114   | 0.343 |
| <b>Knowledge classification (6)</b><br>(KNOWCLASS6) | Fully specialized follower (FSF): $KB_i = 0$ & $KD_i \leq^{geom} \mu_{KD}$            | 95    | 0.286 |
|   | Fully specialized expert (FSE): $KB_i = 0$ & $KD_i >^{geom} \mu_{KD}$                 | 34    | 0.102 |
|   | Mildly diversified follower (MDF): $0 < KB_i \leq 0.77$ & $KD_i \leq^{geom} \mu_{KD}$ | 54    | 0.163 |
|   | Mildly diversified expert (MDE): $0 < KB_i \leq 0.77$ & $KD_i >^{geom} \mu_{KD}$      | 73    | 0.220 |
|   | Highly diversified follower (HDF): $KB_i > 0.77$ & $KD_i \leq^{geom} \mu_{KD}$        | 3     | 0.009 |
|   | Highly diversified expert (HDE): $KB_i > 0.77$ & $KD_i >^{geom} \mu_{KD}$             | 73    | 0.220 |

Note:  $M_{KB} = \text{median}(KB) = 0.4821$  ;  $M_{KD} = \text{median}(\ln(KD)) = 5.9879e-004 = e^{-7.42}$  ;  $^{geom} \mu_{KD} = 5.2090e-004 = e^{-7.56}$

**Table 23: Correlation between independent variables and collinearity diagnostic**

|  | KB     | KD     | ln(KD) | PC | Variable   | VIF  | VIF <sup>2</sup> | Tolerance | R <sup>2</sup> |
|--|--------|--------|--------|----|------------|------|------------------|-----------|----------------|
| KB   | 1      |        |        |    | KB         | 1.82 | 1.35             | 0.5505    | 0.4495         |
| KD   | 0.3003 | 1      |        |    | ln(KD)     | 1.66 | 1.29             | 0.6023    | 0.3977         |
| ln(KD)   | 0.5559 | 0.5718 | 1      |    | PC         | 1.72 | 1.31             | 0.5807    | 0.4193         |
| PC   | 0.5894 | 0.5176 | 0.5387 | 1  | Experience | 1.09 | 1.04             | 0.9206    | 0.0794         |
| <i>All correlation coefficients have p-value=0.000</i> |        |        |        |    | Mean VIF   | 1.57 |                  |           |                |



**Figure 34: Scatter plot of continuous variables**

As a preliminary investigation of the univariate effects of our predictors on the probability of innovative survival, we have also computed the empirical conditional probability of survival for core innovators, given their values of KB, KD, PC, KNOWCLASS, experience and business type and the level of knowledge persistence. These are reported in the Appendix A.5.2. They show that diversification and knowledge accumulation are associated with better survival chances but only after a given threshold is passed. Having large values of both KB and KD as well as increasing the size of the technological knowledge base is associated with a larger probability of survival. We have also assessed whether survivors and exiters can be legitimately be described as belonging to two different populations. We did that by testing whether the empirical cumulative distribution function of KB and KD for survivors is statistically dominated by the one for exiters. These tests are reported in the Appendix A.5.3. They show that indeed survivors are statistically more likely to have larger KB and KD than exiters. We shall see whether putting all these possible explanations of survival together into econometric regressions will confirm these preliminary insights.

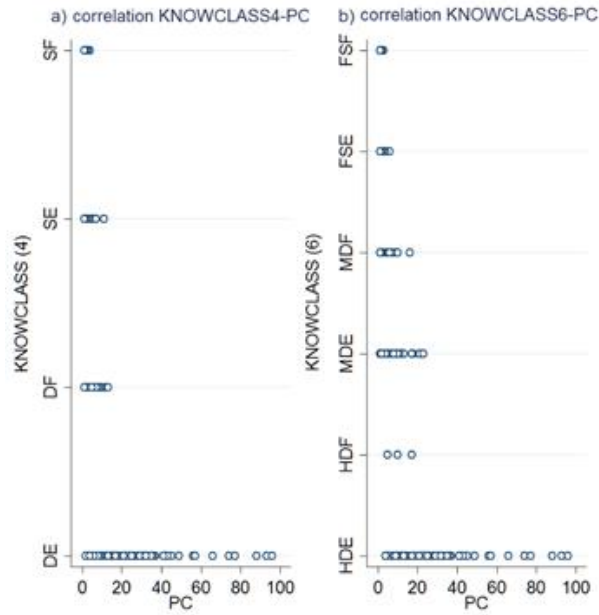


Figure 35: Scatter plot of categorical variables versus technical knowledge base size

## 5.8 Findings

### *Hypothesis 1: Knowledge diversification*

Table 24 reports the estimated coefficients for a set of logit models. We start from the simplest possible model in which we estimate the linear effect of diversification and knowledge accumulation without controlling for any other effect. This returns the coefficients reported in Model 1. We then progressively include control variables to assess how the sign and strength of the coefficients of KB and KD change when we start controlling for alternative explanations. Once we have assessed the effect of the control variables, we add the last variable of interest, the level of knowledge persistence from period to period. This lead to the full model sketched in Section 5.4 and reported in Table 24 as Model 5. This exercise confirms that knowledge diversification significantly affects the probability of survival. However, this is statistically true only for large levels of KB. Bins higher than the sixth one have higher and statistically significant better odds of survival than fully specialized players (the baseline is the first bin in which there are only observations with KB=0). As we discussed in the previous section, we refrain to conclude anything about the relationship between diversification and innovative survival in the range  $0 < KB \leq 0.4$ . However, we note that core innovators with a level of KB in the range  $0.4 < KB \leq 0.6$  (i.e. bins 5 and 6) do not enjoy better survival chances than fully specialized players. Bins 5 and 6 are not reported to save space but their coefficient is not significant. The positive effect of technical knowledge diversification on innovative survival, emerges from the seventh bin of KB. The coefficients reported in the tables are the slopes of the relationships. In a logit model, they can be easily related to the odds of survival. After controlling for technological size, business model, experience, and level of knowledge persistence, we estimate that, compared to the first KB bin (i.e. KB=0), the estimated odds of survival are  $e^{1.734}=5.66$  times higher for the seventh bin,  $e^{2.905}=18.26$  higher for the eighth bin and  $e^{1.740}=5.66$  for the ninth one. The tenth bin had to be excluded from the estimation because it perfectly predicted survival (i.e. all observations with  $KB \geq 0.9$  survived to the next period). These effects are all statistically significant at the 95% confidence level. The coefficients for Model 7, where we assumed a relationship between the log-



odds of survival and KB defined by a restricted cubic spline function, confirm the significant effect of KB. The joint effect of the restricted cubic spline coefficients is significant at the 95% level. However, the first coefficient is not. This suggests that this function does not perfectly capture the effect of KB. Most likely this is due to the scarcity of observation for low levels of knowledge diversification.

Because of this reason, it is particularly insightful to look at the estimated adjusted predicted probability of survival and marginal effects of diversification. These are shown in Figure 36. The former is the estimated probability of survival conditional to the level of KB only, i.e. filtered by the estimated effect of the other variables included in the regression. The latter are the first derivative of the former, i.e. they indicate how much survival chances improve for a decimal unit change in KB. We compute adjusted probabilities and average marginal effects using observed values for the variables other than the one of interest. This means that for each observation we fix the variable of interest at a given value (e.g. we increment KB from zero to one of one decimal point at the time), keep the other variables at their observed value and compute the predicted probability of innovative survival using the coefficients estimated by the regression. Finally, we average the predicted values<sup>24</sup>. This is preferable than using the means values of the other predictors as the observed values allow estimating the effect of the variable of interests more precisely by replicating the economic and statistical logic of matching studies in which *all else is equal*. In Panels A and C and E of Figure 36, we also overlay the scatter plot for the predicted probability of survival estimated for each observation including all possible variables. The comparison with the adjusted probability of survival for KB helps identifying what is the effect of KB on the overall probability of survival, given the combined effects of all the other variables. Panel A reports the estimations for Model 9 in Table 24. The conclusion emerged from the regression coefficients is confirmed. The probability of survival for diversified players is significantly larger than for fully specialized ones but only from the seventh bin onwards. In particular, the adjusted probability of survival is 0.65 for fully specialized innovators and increases to more than 0.89 for players with  $KB \geq 0.6$ . From this level on further increasing KB statistically improves survival chances, as confirmed by the estimated average marginal effects. Finally, we look at adjusted probability of survival and marginal effects for Model 10 in Table 24. These are shown in Panel C of Figure 36. The use of a restricted cubic spline function of KB, seems to let a U-shaped relationship for the adjusted probability of survival to emerge. However, at no levels of KB the estimated  $P(S)$  for diversified players is statistically different from that of fully specialized ones. Yet, the marginal effects confirms the findings from Model 9, namely that starting from intermediate levels of KB further diversifying increases the probability of survival, although with decreasing returns.

We can derive the following conclusions out of the combined analysis of the estimated coefficients for KB and the adjusted probabilities and marginal effects for all models. First, diversification statistically increases survival chances only from intermediate levels of KB onwards and it seems to so with decreasing returns. Second, we cannot state any safe conclusion for what happens at lower levels of KB because the estimation is very inefficient in the range of values  $0 < KB \leq 0.4$  and particularly poor for KB equal to zero. Indeed, in all our models the predicted probability of survival for fully specialized players has a very high variance (as indicated by the high range of value taken by the scatter plots). Moreover, the adjusted probability of survival for

---

<sup>24</sup> The procedure uses the STATA margins command as explained in Williams (2012).

fully specialized players greatly varies when we shift from a categorical variable of KB to a restricted cubic spline function. It fluctuates from 0.65 to 0.95 depending on the model estimated. This, together with the scarcity of observations in the subsequent range  $0 < KB \leq 0.4$ , might strongly affect the shape of the relationship for the restricted cubic spline model, whose U-shape could consequently just be an artefact.

In this respect, it is also worth highlighting that the coefficient for the *Specialized* dummy, which we add in the restricted cubic spline model to estimate the probability of survival for KB equal to zero, is not significant. This points to a lack of a peculiar effect of being fully specialized compared to being diversified at any level. This, together with the high variance of predicted  $P(S)$  for KB equal to zero, points to possibly idiosyncratic explanations of survival for fully specialized innovators. To control for the presence of firm-specific effects and the possibility that survival of fully specialized innovators follows different determinants than for diversified ones, we perform two distinct exercises. First, we break the sample into two sub-samples, one for  $KB=0$  and one for  $KB>0$ . Second, we run the same set of regressions estimating a panel data random effect logistic regression model.

Table 25 reports the coefficients of the models estimated using the two sub-samples for fully specialized and diversified core innovators. We comment on the first two models in the next sections, as there is obviously no diversification effect for the subsample of observations with KB equal to zero. Model 3 in Table 25 replicates Models 6 reported in Table 24. The same conclusions emerged from the latter hold, namely that diversification significantly improves survival chances when innovators strongly diversify. Note that in Model 3 of Table 25 the baseline is now bin 5, i.e. the first one that we can trust in the subsample of diversified players due to the scarcity of observations in bins 2, 3 and 4. This strongly indicates that our conclusions do not depend on the estimation of survival chances for fully specialized players.

Models 8 and 9 in Table 24 and 4, 5 and 6 in Table 25, report the panel data random effect logit version of the pooled data models discussed thus far. Results are broadly consistent with what emerged from the analysis of the pooled data with clustered errors. Coefficients for the KB variables are higher for KB bin 7 and 8 in the random effect panel data Model 8 than in the corresponding pooled data one (i.e. Model 6). However, the coefficient for Bin 9 loses significance. This suggests that the high survival chances enjoyed by fully diversified players are idiosyncratic to their capabilities and not related to their high level of KB.

The contribution of the panel level variance component  $\sigma_u$  is highly significant. In both cases (Models 8 and 9),  $\rho$  is statistically different from zero and equal to 0.6. This means that around sixty percent of the variance in the probability of survival is explained by intra-cluster correlation between the probabilities of survival for a given player in different periods. As explained by Rodriguez and Elo (2003, p.43), “the estimate of  $\sigma_u$  can be interpreted as an ordinary logit coefficient by writing the random effect  $u_{ij} \sim N(0, \sigma_u^2)$  as  $\sigma_u z_{ij}$ , where  $z_{ij} \sim N(0, 1)$ . In this formulation, there is a parallel between the covariates  $x_{ij}$ , representing observed characteristics with coefficients  $\beta$ , and the standardized random effects  $z_{ij}$ , representing unobserved traits with coefficient  $\sigma_u$ ”. More precisely, if we take Model 8 as an example, the odds of survival in a given year for a core innovator who has unobserved propensity to survive one standard deviation above the mean are almost nine times the corresponding odds for a core innovator with average unobserved propensity and the same observed characteristics ( $e^{\sigma_u} = e^{2.185} = 8.89$ ).

For the sake of comparison, we also computed adjusted predictions and marginal effects for the fixed-effect panel data model where we use the categorical variable for KB. Both are computed by setting the random intercept (i.e. the intra-class panel variance component  $\mu_j$ ) to zero. In other words, they show the effect of diversification after we filter the unobserved play-

er-specific effect out. They are shown in Panel E and F of Figure 36. They confirm the main conclusions discussed thus far. Technical knowledge diversification improves innovative survival chances from intermediate levels of KB onwards, even after controlling for firm-specific random effects. The figures also show that the random-effect panel data models failed to reduce the variance of predictions for fully specialized players. We perform the same exercise of splitting the sample done for the pooled data to see if random effects differently affect fully specialized and diversified players. The estimated coefficients are reported Models 4, 5 and 6 in Table 25. An interesting finding emerges. In principle, the inclusion of random effects wash out the significance of diversification for the  $KB > 0$  sample. The strength of the effect actually increases (the coefficients are larger) but it loses significance. This might be easily interpreted as a sign that for diversified players the firm-specific effect is the actual determinant of survival that is masked as diversification effect when we neglect to consider both. However, the panel estimator is not significantly different from zero. This means that for the split sample, the two random-effect panel data models are not better than the pooled data versions discussed before.

To sum up, considering the commonalities of findings about the effects of knowledge breadth on innovative survival we can conclude that starting from intermediate levels of knowledge breadth, further diversification significantly increases the resilience of core innovators and their ability to survive. This emerged from the comprehensive analysis of estimated coefficients, adjusted probabilities, marginal effects, controlling for player-specific random effects and splitting the sample in fully specialized and diversified players. We therefore confirm our first hypothesis within these constraints. Due to the statistical characteristics of our sample, the peculiar distribution of KB and the lack of more detailed information at the company level, we are unable to derive any safe conclusion about the effect of diversification when players have small levels of diversification.

### *Hypothesis 2: Knowledge accumulation*

Our second hypothesis focuses on the role of knowledge accumulation. If we look at Table 24, we notice that the effect of KD on the log-odds of survival is positive and significant as long as we do not control for experience and knowledge persistence. Both variables are related to time. Not surprisingly, this suggests that knowledge accumulation is at least partially correlated with time. In particular, the lack of significance of KD when we control for the time-related variables means that the informative content included in the former is actually just a manifestation of the latter. Table 25 provides the same findings regardless of the composition of the sample. The lack of significance for the coefficients of KD, no matter if we assume a linear or quadratic relationship between KD and the odds of innovative survival or if we look at the entire population or at specialized and diversified innovators separately, lead us to reject hypothesis two.

**Table 24: Pooled and panel data logit model estimation**

| LOGIT  | (1)       | (2)       | (3)       | (4)       | (5)       | (6)       | (7)       | (8)      | (9)      |
|--|-----------|-----------|-----------|-----------|-----------|-----------|-----------|----------|----------|
| MODEL  | Pooled    | Pooled    | Pooled    | Pooled    | Pooled    | Pooled    | Pooled    | Panel    | Panel    |
| VARIABLES                                    | clust.VCE | clust.VCE | clust.VCE | clust.VCE | clust.VCE | clust.VCE | clust.VCE | RandEff  | RandEff  |
|  | PC<24     | PC<24     | PC<24     | PC<24     | PC<24     | PC<24     | All obs.  | PC<24    | All obs. |
|  | survival  | survival  | survival  | survival  | survival  | survival  | survival  | survival | survival |
| <b>KB bin category (base = bin 1 = KB=0)</b> |           |           |           |           |           |           |           |          |          |
| KB bin 7                                     | 1.398**   | 1.615**   | 1.715**   | 1.808***  | 1.719**   | 1.734**   |           | 2.361**  |          |
| (0.6≤KB<0.7)                                 | (0.630)   | (0.653)   | (0.691)   | (0.688)   | (0.688)   | (0.705)   |           | (1.198)  |          |
| KB bin 8                                     | 2.679**   | 3.159**   | 3.207**   | 3.163**   | 2.925**   | 2.905**   |           | 4.244**  |          |
| (0.7≤KB<0.8)                                 | (1.044)   | (1.331)   | (1.286)   | (1.317)   | (1.346)   | (1.347)   |           | (1.928)  |          |

| LOGIT<br>MODEL<br>VARIABLES   | (1)<br>Pooled<br>clust.VCE<br>PC<24<br>survival | (2)<br>Pooled<br>clust.VCE<br>PC<24<br>survival | (3)<br>Pooled<br>clust.VCE<br>PC<24<br>survival | (4)<br>Pooled<br>clust.VCE<br>PC<24<br>survival | (5)<br>Pooled<br>clust.VCE<br>PC<24<br>survival | (6)<br>Pooled<br>clust.VCE<br>PC<24<br>survival | (7)<br>Pooled<br>clust.VCE<br>All obs.<br>survival | (8)<br>Panel<br>RandEff<br>PC<24<br>survival | (9)<br>Panel<br>RandEff<br>All obs.<br>survival |
|---|---|---|---|---|---|---|--|--|---|
| KB bin 9<br>(0.8≤KB<0.9)  | 1.392*<br>(0.800)                               | 2.130***<br>(0.739)                             | 2.032***<br>(0.732)                             | 2.049***<br>(0.756)                             | 1.725**<br>(0.761)                              | 1.740**<br>(0.759)                              |  | 3.093<br>(2.029)                             |   |
| <b>Specialized</b>  |   |   |   |   |   |   | -2.142<br>(1.783)                                  |  | -3.853<br>(3.231)                               |
| <b>Restricted cubic spline</b> ( <i>knots at KB = 0, 0.48, 0.89</i> ) |   |   |   |   |   |   |  |  |   |
| <b>r.c.s. KB1</b>   |   |   |   |   |   |   | -7.273<br>(5.521)                                  |  | -13.52<br>(9.904)                               |
| <b>r.c.s. KB2</b>   |   |   |   |   |   |   | 10.73*<br>(5.701)                                  |  | 18.46*<br>(10.38)                               |
| <b>ln(KD)</b>   | 0.148*<br>(0.0827)                              | 0.166*<br>(0.0877)                              | 0.128<br>(0.0914)                               | 0.141<br>(0.0981)                               | 0.0974<br>(0.100)                               | -0.423<br>(0.443)                               | -0.390<br>(0.447)                                  | -0.855<br>(0.683)                            | -0.822<br>(0.670)                               |
| <b>ln(KD)<sup>2</sup></b>   |   |   |   |   |   | -0.0295<br>(0.0251)                             | -0.0280<br>(0.0253)                                | -0.0601<br>(0.0402)                          | -0.0589<br>(0.0395)                             |
| <b>PC</b>   |   | -0.0626<br>(0.0639)                             | -0.0262<br>(0.0687)                             | -0.0355<br>(0.0715)                             | -0.0146<br>(0.0721)                             | -0.00655<br>(0.0729)                            | 0.0124<br>(0.0351)                                 | -0.0465<br>(0.103)                           | 0.00340<br>(0.0624)                             |
| <b>Experience</b> ( <i>base = new entrants</i> )                      |   |   |   |   |   |   |  |  |   |
| 1 period  |   |   | -1.367***<br>(0.367)                            | -1.320***<br>(0.395)                            | -1.340***<br>(0.411)                            | -1.345***<br>(0.410)                            | -1.271***<br>(0.408)                               | -2.208**<br>(0.964)                          | -2.169**<br>(0.963)                             |
| 2 periods   |   |   | -1.302***<br>(0.456)                            | -1.279***<br>(0.482)                            | -1.118**<br>(0.486)                             | -1.113**<br>(0.489)                             | -1.039**<br>(0.477)                                | -2.401*<br>(1.295)                           | -2.433*<br>(1.325)                              |
| 3 periods   |   |   | -0.900<br>(0.593)                               | -0.956<br>(0.664)                               | -0.531<br>(0.677)                               | -0.426<br>(0.637)                               | -0.457<br>(0.621)                                  | -1.793<br>(1.479)                            | -1.877<br>(1.497)                               |
| 4 periods   |   |   | -1.161*<br>(0.602)                              | -1.175*<br>(0.692)                              | -0.704<br>(0.699)                               | -0.707<br>(0.711)                               | -0.612<br>(0.693)                                  | -2.493<br>(1.807)                            | -2.324<br>(1.738)                               |
| <b>Business cat.</b> ( <i>base = IDMs</i> )                           |   |   |   |   |   |   |  |  |   |
| Fabless   |   |   |   | -0.579<br>(0.535)                               | -0.298<br>(0.557)                               | -0.285<br>(0.575)                               | -0.145<br>(0.583)                                  | -0.915<br>(1.429)                            | -0.693<br>(1.417)                               |
| Supplier  |   |   |   | 0.628<br>(0.852)                                | 0.987<br>(0.837)                                | 0.967<br>(0.853)                                | 0.949<br>(0.818)                                   | 1.291<br>(1.387)                             | 1.379<br>(1.396)                                |
| Res.provider  |   |   |   | -0.832*<br>(0.472)                              | -0.686<br>(0.499)                               | -0.697<br>(0.507)                               | -0.727<br>(0.486)                                  | -1.378<br>(0.944)                            | -1.321<br>(0.929)                               |
| User  |   |   |   | -0.429<br>(0.426)                               | -0.427<br>(0.424)                               | -0.405<br>(0.423)                               | -0.543<br>(0.412)                                  | -0.856<br>(0.861)                            | -0.980<br>(0.862)                               |
| <b>RK Persistence</b><br>( <i>main traject.</i> )                     |   |   |   |   | 18.19**<br>(7.919)                              | 18.11**<br>(7.811)                              | 18.15**<br>(7.625)                                 | 27.13*<br>(15.04)                            | 28.73*<br>(15.15)                               |
| <b>Constant</b>   | 1.882**<br>(0.777)                              | 2.132**<br>(0.861)                              | 2.646***<br>(0.946)                             | 2.977***<br>(0.999)                             | 0.559<br>(1.442)                                | -1.603<br>(2.214)                               | 0.634<br>(2.549)                                   | -2.465<br>(3.323)                            | 1.331<br>(4.336)                                |
| Observations  | 290   | 290   | 290   | 286   | 286   | 286   | 323  | 286  | 323   |
| # of companies  |   |   |   |   |   |   |  | 122  | 126   |
| Log-Likelihood  | -145.9  | -145.4  | -138.6  | -135.3  | -132.8  | -132.1  | -135.6   | -129.8                                       | -133.1  |
| Pseudo R2   | 0.114   | 0.117   | 0.158   | 0.173   | 0.188   | 0.192   | 0.220  |  |   |
| Sigma_u   |   |   |   |   |   |   |  | 2.185  | 2.273   |
| Rho   |   |   |   |   |   |   |  | 0.592  | 0.611   |
| Signif. of rho  |   |   |   |   |   |   |  | 0.016  | 0.013   |

**Notes:** Robust standard errors in parentheses (errors clustered by company for the pooled data)

Significance legend: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Vertical bars indicate joint significance of the variables reported to the left of the bar(s) (||| p<0.01, || p<0.05, | p<0.1). Business category 'Foundry' (in all models) and KB bins 2 and 10 (in Model 9) were omitted from the regressions as they predicted survival perfectly. For Model 9 only KB bins whose coefficient is statistically significant are reported.

**Table 25: Pooled and panel data logit model estimation with split sample**

| LOGIT                                   | (1)       | (2)       | (3)          | (4)       | (5)       | (6)          |
|---|-----------|-----------|--------------|-----------|-----------|--------------|
| MODEL                                   | Pooled    | Pooled    | Pooled       | Panel     | Panel     | Panel        |
| VARIABLES                               | clust.VCE | clust.VCE | clust.VCE    | RandEff   | RandEff   | RandEff      |
|   | KB=0      | KB=0      | PC<24 & KB>0 | KB=0      | KB=0      | PC<24 & KB>0 |
|   | survival  | survival  | survival     | survival  | survival  | survival     |
| <b>KB bin category (base = bin 5)</b>   |           |           |              |           |           |              |
| KB bin 6                                |           |           | 0.660        |           |           | 0.786        |
| (0.5≤KB<0.6)                            |           |           | (0.585)      |           |           | (0.982)      |
| KB bin 7                                |           |           | 1.926**      |           |           | 2.460        |
| (0.6≤KB<0.7)                            |           |           | (0.902)      |           |           | (1.650)      |
| KB bin 8                                |           |           | 3.286**      |           |           | 4.631        |
| (0.7≤KB<0.8)                            |           |           | (1.293)      |           |           | (3.330)      |
| KB bin 9                                |           |           | 2.419**      |           |           | 3.530        |
| (0.8≤KB<0.9)                            |           |           | (0.968)      |           |           | (2.897)      |
| <b>ln(KD)</b>                           | 0.0311    | -0.819    | 1.793        | 0.0311    | -0.819    | 2.086        |
|   | (0.109)   | (0.508)   | (1.411)      | (0.110)   | (0.631)   | (1.879)      |
| <b>ln(KD)<sup>2</sup></b>               |           | -0.0466*  | 0.0940       |           | -0.0466   | 0.103        |
|   |           | (0.0283)  | (0.0812)     |           | (0.0339)  | (0.108)      |
| <b>PC</b>                               | 0.0523    | 0.0925    | -0.0990      | 0.0525    | 0.0925    | -0.158       |
|   | (0.260)   | (0.249)   | (0.0714)     | (0.261)   | (0.265)   | (0.163)      |
| <b>Experience (base = new entrants)</b> |           |           |              |           |           |              |
| 1 period                                | -1.711**  | -1.730**  | -0.995       | -1.710**  | -1.730**  | -1.598       |
|   | (0.699)   | (0.706)   | (0.729)      | (0.684)   | (0.685)   | (1.746)      |
| 2 periods                               | -1.921**  | -1.907**  | -0.224       | -1.920*** | -1.907*** | -0.590       |
|   | (0.769)   | (0.792)   | (0.801)      | (0.728)   | (0.736)   | (1.485)      |
| 3 periods                               | -1.965*   | -1.674*   | 1.559        | -1.965**  | -1.674*   | 1.735        |
|   | (1.005)   | (0.994)   | (1.337)      | (0.921)   | (0.951)   | (1.803)      |
| 4 periods                               | -1.586    | -1.656    | -0.0383      | -1.585    | -1.656    | -0.539       |
|   | (1.024)   | (1.024)   | (1.163)      | (1.012)   | (1.023)   | (2.132)      |
| <b>Business cat. (base = IDMs)</b>      |           |           |              |           |           |              |
| Fabless                                 | -0.679    | -0.683    | -            | -0.679    | -0.683    | -            |
|   | (0.789)   | (0.796)   |              | (0.874)   | (0.871)   |              |
| Supplier                                | 0.550     | 0.513     | 0.712        | 0.550     | 0.513     | 1.164        |
|   | (1.351)   | (1.376)   | (1.217)      | (1.261)   | (1.268)   | (2.201)      |
| Res.provider                            | -0.0841   | -0.0752   | -1.304**     | -0.0841   | -0.0752   | -2.004       |
|   | (0.603)   | (0.628)   | (0.657)      | (0.611)   | (0.620)   | (1.872)      |
| User                                    | -0.673    | -0.647    | -0.812       | -0.673    | -0.647    | -1.308       |
|   | (0.570)   | (0.559)   | (0.644)      | (0.626)   | (0.621)   | (1.460)      |
| <b>RK Persistence</b>                   | 40.61***  | 39.32***  | 0.224        | 40.60***  | 39.32***  | 4.782        |
| (main traject.)                         | (13.02)   | (12.55)   | (12.14)      | (14.62)   | (14.41)   | (21.71)      |
| <b>Constant</b>                         | -1.787    | -5.399**  | 9.586        | -1.787    | -5.399    | 11.82        |
|   | (1.869)   | (2.752)   | (6.371)      | (2.149)   | (3.462)   | (9.515)      |
| Observations                            | 128       | 128       | 155          | 128       | 128       | 155          |
| # of companies                          |           |           |              | 84        | 84        | 83           |
| Log-Likelihood                          | -66.36    | -65.32    | -54.71       | -66.36    | -65.32    | -54.53       |
| Pseudo R2                               | 0.211     | 0.224     | 0.237        |           |           |              |
| σ <sub>u</sub>                          |           |           |              | 0.0027    | 0.000978  | 1.866        |
| Rho                                     |           |           |              | 2.34e-06  | 2.91e-07  | 0.514        |
| Significance of rho                     |           |           |              | 0.498     | 0.499     | 0.273        |

**Notes:** Robust standard errors in parentheses (errors clustered by company for the pooled data)

Significance legend: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Vertical bars indicates joint significance (||| p<0.01, || p<0.05, | p<0.1).

Business category 'Foundry' (in all models), 'Fabless' and KB bins 2 and 10 (in Model 3 and 6) are omitted from the regressions as they predict survival perfectly. For Models 3 and 6 only KB bins others than 2, 3 and 4 are reported. Bin is 5 used as baseline.

**Table 26: Pooled data logit models estimation with KNOWCLASS**

| LOGIT<br>MODEL<br>VARIABLE              | (1)<br>PC<15<br>survival | (2)<br>PC<15<br>survival | (3)<br>PC<15<br>survival | (4)<br>PC<15<br>survival | (5)<br>PC<24<br>survival | (6)<br>PC<24<br>survival | (7)<br>PC<24<br>survival | (8)<br>PC<24<br>survival | (9)<br>PC<24<br>survival | (10)<br>PC<24<br>survival |
|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|---------------------------|
| <b>Knowclass (4):</b>                   |                          |                          |                          |                          |                          |                          |                          |                          |                          |                           |
| Div. Expert                             | 2.43***<br>(0.872)       | 2.158**<br>(0.853)       | 1.836**<br>(0.813)       |                          |                          |                          |                          |                          |                          |                           |
| Div. Follower                           | 0.597<br>(0.507)         | 0.322<br>(0.462)         |                          | -1.83**<br>(0.813)       |                          |                          |                          |                          |                          |                           |
| Spec. Expert                            | 0.275<br>(0.397)         |                          | -0.322<br>(0.462)        | -2.15**<br>(0.853)       |                          |                          |                          |                          |                          |                           |
| Spec. Follower                          |                          | -0.275<br>(0.397)        | -0.597<br>(0.507)        | -2.4***<br>(0.872)       |                          |                          |                          |                          |                          |                           |
| <b>Knowclass (6):</b>                   |                          |                          |                          |                          |                          |                          |                          |                          |                          |                           |
| Highly div. expert                      |                          |                          |                          |                          | 1.529**<br>(0.612)       | 1.271*<br>(0.667)        | 1.105*<br>(0.623)        | 0.406<br>(0.592)         | 2.780**<br>(1.373)       |                           |
| Highly div. follow                      |                          |                          |                          |                          | -1.250<br>(1.483)        | -1.508<br>(1.485)        | -1.674<br>(1.386)        | -2.374*<br>(1.342)       |                          | -2.78**<br>(1.373)        |
| Inter. div. expert                      |                          |                          |                          |                          | 1.124**<br>(0.548)       | 0.866<br>(0.629)         | 0.700<br>(0.543)         |                          | 2.374*<br>(1.342)        | -0.406<br>(0.592)         |
| Inter. div. follow.                     |                          |                          |                          |                          | 0.424<br>(0.431)         | 0.166<br>(0.497)         |                          | -0.700<br>(0.543)        | 1.674<br>(1.386)         | -1.105*<br>(0.623)        |
| Spec. expert                            |                          |                          |                          |                          | 0.258<br>(0.485)         |                          | -0.166<br>(0.497)        | -0.866<br>(0.629)        | 1.508<br>(1.485)         | -1.271*<br>(0.667)        |
| Spec. follower                          |                          |                          |                          |                          |                          | -0.258<br>(0.485)        | -0.424<br>(0.431)        | -1.12**<br>(0.548)       | 1.250<br>(1.483)         | -1.52**<br>(0.612)        |
| <b>PC</b>                               | 0.120<br>(0.098)         | 0.120<br>(0.098)         | 0.120<br>(0.098)         | 0.120<br>(0.098)         | 0.0826<br>(0.085)        | 0.0826<br>(0.085)        | 0.0826<br>(0.085)        | 0.0826<br>(0.085)        | 0.0826<br>(0.085)        | 0.0826<br>(0.085)         |
| <b>Experience</b>                       |                          |                          |                          |                          |                          |                          |                          |                          |                          |                           |
| 1 period                                | -1.2***<br>(0.446)       | -1.2***<br>(0.446)       | -1.2***<br>(0.446)       | -1.2***<br>(0.446)       | -1.2***<br>(0.427)       | -1.2***<br>(0.427)       | -1.2***<br>(0.427)       | -1.2***<br>(0.427)       | -1.2***<br>(0.427)       | -1.2***<br>(0.427)        |
| 2 periods                               | -1.19**<br>(0.488)       | -1.19**<br>(0.488)       | -1.19**<br>(0.488)       | -1.19**<br>(0.488)       | -1.03**<br>(0.487)       | -1.03**<br>(0.487)       | -1.03**<br>(0.487)       | -1.03**<br>(0.487)       | -1.03**<br>(0.487)       | -1.03**<br>(0.487)        |
| 3 periods                               | -0.679<br>(0.642)        | -0.679<br>(0.642)        | -0.679<br>(0.642)        | -0.679<br>(0.642)        | -0.532<br>(0.664)        | -0.532<br>(0.664)        | -0.532<br>(0.664)        | -0.532<br>(0.664)        | -0.532<br>(0.664)        | -0.532<br>(0.664)         |
| 4 periods                               | -0.802<br>(0.708)        | -0.802<br>(0.708)        | -0.802<br>(0.708)        | -0.802<br>(0.708)        | -0.350<br>(0.700)        | -0.350<br>(0.700)        | -0.350<br>(0.700)        | -0.350<br>(0.700)        | -0.350<br>(0.700)        | -0.350<br>(0.700)         |
| <b>Business cat.</b>                    |                          |                          |                          |                          |                          |                          |                          |                          |                          |                           |
| Foundry                                 | -                        | -                        | -                        | -                        | -                        | -                        | -                        | -                        | -                        | -                         |
| Fabless                                 | -0.307<br>(0.559)        | -0.307<br>(0.559)        | -0.307<br>(0.559)        | -0.307<br>(0.559)        | -0.147<br>(0.552)        | -0.147<br>(0.552)        | -0.147<br>(0.552)        | -0.147<br>(0.552)        | -0.147<br>(0.552)        | -0.147<br>(0.552)         |
| Supplier                                | 0.907<br>(0.806)         | 0.907<br>(0.806)         | 0.907<br>(0.806)         | 0.907<br>(0.806)         | 1.068<br>(0.817)         | 1.068<br>(0.817)         | 1.068<br>(0.817)         | 1.068<br>(0.817)         | 1.068<br>(0.817)         | 1.068<br>(0.817)          |
| Res. provider                           | -0.750<br>(0.488)        | -0.750<br>(0.488)        | -0.750<br>(0.488)        | -0.750<br>(0.488)        | -0.689<br>(0.465)        | -0.689<br>(0.465)        | -0.689<br>(0.465)        | -0.689<br>(0.465)        | -0.689<br>(0.465)        | -0.689<br>(0.465)         |
| User                                    | -0.832*<br>(0.454)       | -0.832*<br>(0.454)       | -0.832*<br>(0.454)       | -0.832*<br>(0.454)       | -0.766*<br>(0.411)       | -0.766*<br>(0.411)       | -0.766*<br>(0.411)       | -0.766*<br>(0.411)       | -0.766*<br>(0.411)       | -0.766*<br>(0.411)        |
| <b>RK Persistence</b><br>(main trajet.) | 21.79***<br>(7.509)      | 21.79***<br>(7.509)      | 21.79***<br>(7.509)      | 21.79***<br>(7.509)      | 20.54***<br>(7.451)      | 20.54***<br>(7.451)      | 20.54***<br>(7.451)      | 20.54***<br>(7.451)      | 20.54***<br>(7.451)      | 20.54***<br>(7.451)       |
| <b>Constant</b>                         | -0.897<br>(0.875)        | -0.622<br>(0.932)        | -0.299<br>(1.012)        | 1.536<br>(1.231)         | -0.795<br>(0.880)        | -0.537<br>(0.968)        | -0.371<br>(0.942)        | 0.329<br>(1.101)         | -2.045<br>(1.792)        | 0.735<br>(1.099)          |
| Observations                            | 276                      | 276                      | 276                      | 276                      | 292                      | 292                      | 292                      | 292                      | 292                      | 292                       |
| Log-Likelihood                          | -126.7                   | -126.7                   | -126.7                   | -126.7                   | -138.2                   | -138.2                   | -138.2                   | -138.2                   | -138.2                   | -138.2                    |
| Pseudo R2                               | 0.200                    | 0.200                    | 0.200                    | 0.200                    | 0.164                    | 0.164                    | 0.164                    | 0.164                    | 0.164                    | 0.164                     |

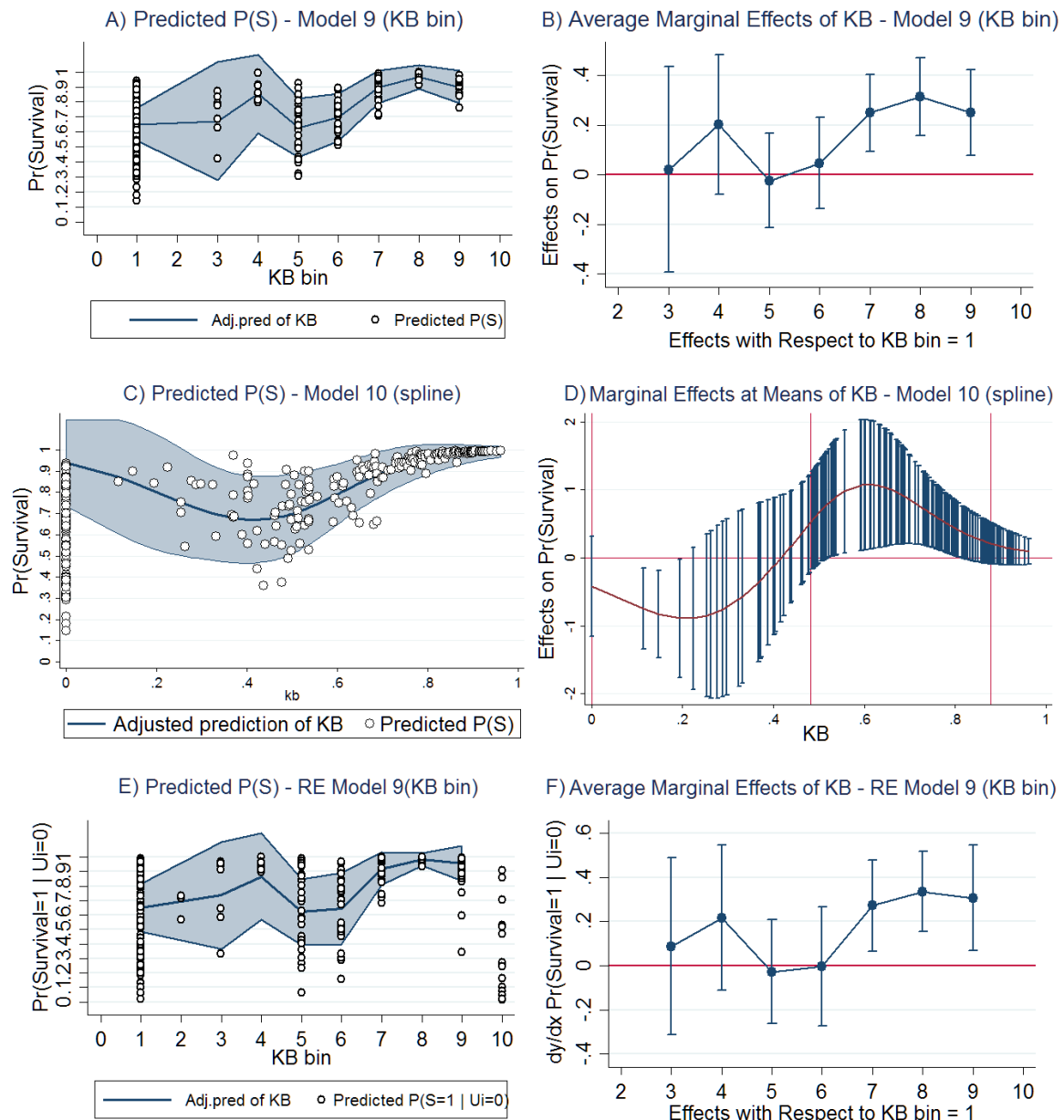
**Notes:** Robust standard errors in parentheses (errors clustered by company). Significance legend: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Business category 'Foundry' (in all models) was omitted from the regressions as it predicts survival perfectly.

**Table 27: Panel data random effect logit models estimation with KNOWCLASS**

| RE XTLOGIT                              | (1)                | (2)                 | (3)                | (4)                 | (5)                | (6)                | (7)                | (8)                | (9)                | (10)               |
|---|--------------------|---------------------|--------------------|---------------------|--------------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| MODEL                                   | PC<15              | PC<15               | PC<15              | PC<15               | PC<24              | PC<24              | PC<24              | PC<24              | PC<24              | PC<24              |
| VARIABLE                                | Surviv.            | Surviv.             | Surviv.            | Surviv.             | Surviv.            | Surviv.            | Surviv.            | Surviv.            | Surviv.            | Surviv.            |
| <b>Knowclass (4):</b>                   |                    |                     |                    |                     |                    |                    |                    |                    |                    |                    |
| Div. Expert                             | 2.93**<br>(1.205)  | 2.48**<br>(1.09)    | 2.13**<br>(1.04)   |                     |                    |                    |                    |                    |                    |                    |
| Div. Follower                           | 0.802<br>(0.617)   | 0.354<br>(0.60)     |                    | -2.132**<br>(1.047) |                    |                    |                    |                    |                    |                    |
| Spec. Expert                            | 0.448<br>(0.556)   |                     | -0.354<br>(0.60)   | -2.485**<br>(1.093) |                    |                    |                    |                    |                    |                    |
| Spec. Follower                          |                    | -0.448<br>(0.55)    | -0.802<br>(0.61)   | -2.934**<br>(1.205) |                    |                    |                    |                    |                    |                    |
| <b>Knowclass (6):</b>                   |                    |                     |                    |                     |                    |                    |                    |                    |                    |                    |
| Highly div. expert                      |                    |                     |                    |                     | 1.846<br>(1.359)   | 1.416<br>(1.335)   | 1.334<br>(1.239)   | 0.537<br>(1.153)   | 2.795<br>(1.911)   |                    |
| Highly div. follow                      |                    |                     |                    |                     | -0.949<br>(1.896)  | -1.379<br>(1.905)  | -1.461<br>(1.827)  | -2.258<br>(1.812)  |                    | -2.795<br>(1.911)  |
| Inter. div. expert                      |                    |                     |                    |                     | 1.309*<br>(0.680)  | 0.879<br>(0.734)   | 0.797<br>(0.626)   |                    | 2.258<br>(1.812)   | -0.537<br>(1.153)  |
| Inter. div. follow.                     |                    |                     |                    |                     | 0.512<br>(0.547)   | 0.0821<br>(0.687)  |                    | -0.797<br>(0.626)  | 1.461<br>(1.827)   | -1.334<br>(1.239)  |
| Spec. expert                            |                    |                     |                    |                     | 0.430<br>(0.639)   |                    | -0.0821<br>(0.687) | -0.879<br>(0.734)  | 1.379<br>(1.905)   | -1.416<br>(1.335)  |
| Spec. follower                          |                    |                     |                    |                     |                    | -0.430<br>(0.639)  | -0.512<br>(0.547)  | -1.309*<br>(0.680) | 0.949<br>(1.896)   | -1.846<br>(1.359)  |
| <b>PC</b>                               | 0.106<br>(0.127)   | 0.106<br>(0.127)    | 0.106<br>(0.127)   | 0.106<br>(0.127)    | 0.0819<br>(0.083)  | 0.0819<br>(0.083)  | 0.0819<br>(0.083)  | 0.0819<br>(0.083)  | 0.0819<br>(0.083)  | 0.0819<br>(0.0831) |
| <b>Experience</b>                       |                    |                     |                    |                     |                    |                    |                    |                    |                    |                    |
| 1 period                                | -1.52**<br>(0.629) | -1.52**<br>(0.629)  | -1.52**<br>(0.629) | -1.52**<br>(0.629)  | -1.53**<br>(0.674) | -1.53**<br>(0.674) | -1.53**<br>(0.674) | -1.53**<br>(0.674) | -1.53**<br>(0.674) | -1.53**<br>(0.674) |
| 2 periods                               | -1.75**<br>(0.866) | -1.752**<br>(0.866) | -1.75**<br>(0.866) | -1.75**<br>(0.866)  | -1.605*<br>(0.937) | -1.605*<br>(0.937) | -1.605*<br>(0.937) | -1.605*<br>(0.937) | -1.605*<br>(0.937) | -1.605*<br>(0.937) |
| 3 periods                               | -1.232<br>(1.020)  | -1.232<br>(1.020)   | -1.232<br>(1.020)  | -1.232<br>(1.020)   | -1.154<br>(1.125)  | -1.154<br>(1.125)  | -1.154<br>(1.125)  | -1.154<br>(1.125)  | -1.154<br>(1.125)  | -1.154<br>(1.125)  |
| 4 periods                               | -1.543<br>(1.219)  | -1.543<br>(1.219)   | -1.543<br>(1.219)  | -1.543<br>(1.219)   | -1.140<br>(1.316)  | -1.140<br>(1.316)  | -1.140<br>(1.316)  | -1.140<br>(1.316)  | -1.140<br>(1.316)  | -1.140<br>(1.316)  |
| <b>Business cat.</b>                    |                    |                     |                    |                     |                    |                    |                    |                    |                    |                    |
| Foundry                                 | -                  | -                   | -                  | -                   | -                  | -                  | -                  | -                  | -                  | -                  |
| Fabless                                 | -0.551<br>(1.021)  | -0.551<br>(1.021)   | -0.551<br>(1.021)  | -0.551<br>(1.021)   | -0.398<br>(1.051)  | -0.398<br>(1.051)  | -0.398<br>(1.051)  | -0.398<br>(1.051)  | -0.398<br>(1.051)  | -0.398<br>(1.051)  |
| Supplier                                | 1.072<br>(1.038)   | 1.072<br>(1.038)    | 1.072<br>(1.038)   | 1.072<br>(1.038)    | 1.270<br>(1.077)   | 1.270<br>(1.077)   | 1.270<br>(1.077)   | 1.270<br>(1.077)   | 1.270<br>(1.077)   | 1.270<br>(1.077)   |
| Res. provider                           | -1.018<br>(0.646)  | -1.018<br>(0.646)   | -1.018<br>(0.646)  | -1.018<br>(0.646)   | -1.036<br>(0.712)  | -1.036<br>(0.712)  | -1.036<br>(0.712)  | -1.036<br>(0.712)  | -1.036<br>(0.712)  | -1.036<br>(0.712)  |
| User                                    | -1.173*<br>(0.658) | -1.173*<br>(0.658)  | -1.173*<br>(0.658) | -1.173*<br>(0.658)  | -1.113<br>(0.687)  | -1.113<br>(0.687)  | -1.113<br>(0.687)  | -1.113<br>(0.687)  | -1.113<br>(0.687)  | -1.113<br>(0.687)  |
| <b>RK Persistence</b><br>(main trajct.) | 27.20**<br>(11.63) | 27.20**<br>(11.63)  | 27.20**<br>(11.63) | 27.20**<br>(11.63)  | 26.67**<br>(12.50) | 26.67**<br>(12.50) | 26.67**<br>(12.50) | 26.67**<br>(12.50) | 26.67**<br>(12.50) | 26.67**<br>(12.50) |
| <b>Constant</b>                         | -0.946<br>(1.278)  | -0.498<br>(1.326)   | -0.144<br>(1.401)  | 1.988<br>(1.713)    | -0.863<br>(1.304)  | -0.433<br>(1.442)  | -0.351<br>(1.416)  | 0.446<br>(1.546)   | -1.812<br>(2.467)  | 0.983<br>(1.945)   |
| Observations                            | 276                | 276                 | 276                | 276                 | 296                | 296                | 296                | 296                | 296                | 296                |
| Num. of innovators                      | 121                | 121                 | 121                | 121                 | 126                | 126                | 126                | 126                | 126                | 126                |
| Log-Likelihood                          | -126.1             | -126.1              | -126.1             | -126.1              | -137.6             | -137.6             | -137.6             | -137.6             | -137.6             | -137.6             |

| RE XTLOGIT          | (1)     | (2)     | (3)     | (4)     | (5)     | (6)     | (7)     | (8)     | (9)     | (10)    |
|---------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| MODEL               | PC<15   | PC<15   | PC<15   | PC<15   | PC<24   | PC<24   | PC<24   | PC<24   | PC<24   | PC<24   |
| VARIAB.             | Surviv. | Surviv. | Surviv. | Surviv. | Surviv. | Surviv. | Surviv. | Surviv. | Surviv. | Surviv. |
| Sigma_u             | 1.218   | 1.218   | 1.218   | 1.218   | 1.371   | 1.371   | 1.371   | 1.371   | 1.371   | 1.371   |
| Rho                 | 0.311   | 0.311   | 0.311   | 0.311   | 0.363   | 0.363   | 0.363   | 0.363   | 0.363   | 0.363   |
| Significance of rho | 0.056   | 0.056   | 0.056   | 0.056   | 0.136   | 0.136   | 0.136   | 0.136   | 0.136   | 0.136   |

**Notes:** Significance legend: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Business category 'Foundry' (in all models) was omitted from the regressions as it predicts survival perfectly.



**Figure 36: Adjusted predictions and average marginal effects of knowledge breadth**

Note: shaded areas and spikes represent 95% confidence intervals. Adjusted probabilities and marginal effects computed at the observed values of the other predictors.



### *Hypothesis 3: Interaction effects between knowledge breadth and accumulation*

To test Hypothesis 3 regarding the presence of complementarities between diversification and knowledge accumulation, we estimate an equivalent version of the full model (Model 5, Table 24) where the individual effects of knowledge breadth and depth are replaced by the categorical variable KNOWCLASS indicating the relative position in the knowledge breadth-depth space. As discussed in Section 5.7, we measure KNOWCLASS in two different ways, corresponding to 4 or 6 categories. Results are reported in Table 26 for the pooled data and in Table 27 for the random effect panel data models. The first four columns report the same model estimated using four knowledge categories based on the median levels of KB and log-KD. The estimated coefficients in each column are computed using each of the four different categories as baseline. Being a diversified expert provides significantly better odds of survival than anyone else. In particular, after controlling for other explanations, including player-specific random effects, diversified experts have  $e^{2.934}=18.8$  times better survival odds than specialized followers. However, moving from any of the other categories to anyone else, does not improve survival chances significantly. This confirms the importance to improve both breadth and depth at the same time, especially for fully specialized players. Yet this is true for the other categories as well. In fact, once you are an expert you can only improve the odds of survival by increasing breadth while maintaining high depth. The same reasoning holds for diversified followers, which need to improve depth while maintaining high breadth to see their odds of survival improving significantly. These conclusions hold true when we measure KNOWCLASS using six categories and pooled data. However, this time controlling for innovator-specific random effects washes significance of the movement across the breadth-depth space out, except for being an intermediate diversified expert compared to a fully specialized follower.

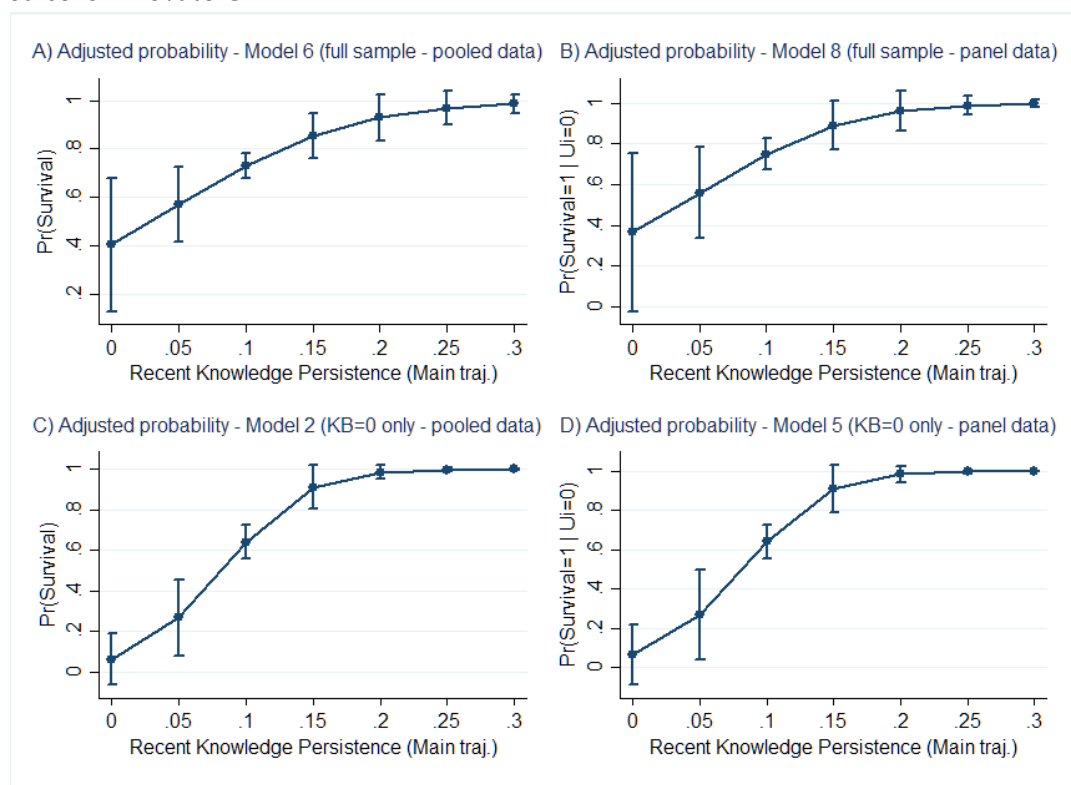
The overall findings provide some evidence in favour of our third hypothesis. Yet we cannot fully confirm it because the results are not robust to the way we categorize the knowledge breadth-depth space, in particular when we control for innovator-specific random effects.

### *Hypothesis 4: Knowledge persistence*

Hypothesis 4 postulated a positive relationship between knowledge persistence and resilience of core innovators. We measured knowledge persistence in several ways, distinguishing between recent and cumulated knowledge and between the main trajectories (identified by the main component of the NMP) and the whole NMP sample of influential patents (including the second component in selected periods). In Tables Table 24, Table 25, Table 26 and Table 27, we only reported recent knowledge persistence in the main technological trajectories. This is the only version of knowledge persistence that turned out significant. This finding is in itself very interesting as it shows that technological shocks that are able to provoke shakeouts come from changes in the approach to problem solving (i.e. the main engineering trajectories identified by the NMPs) concerning the set of problems that currently attracts most of the inventive effort by firms (i.e. the main component of the NMPs).

For the full sample, the variable measuring recent knowledge persistence on the main trajectory is always strongly significant, regardless of the model estimated. The coefficients across the different models show that full recent knowledge persistence on the main trajectory increases log-odds of survival by a value within the range 18.11 – 28.73, compared to zero persistence. As plotted in panel A of Figure 37, this translates in an increase in the probability of survival from 0.4 to 1. When we split the sample between fully specialized and diversified innovators, an interesting finding emerges. Recent knowledge persistence on the main trajectory significantly improves survival chances for the former only whereas it is not significant for the latter. All else

equal and after filtering player-specific random effects out, for fully specialized innovators a change from zero recent knowledge persistence on the main trajectory to full persistence improves the probability of survival from 0.06 to 1, as shown in Panel D. We note from Panels B and C of Figure 37 that a level of about 0.2 is sufficient to achieve a probability of survival of 1 for fully specialized innovators. This allows quantifying a guiding threshold that indicates the level of change in technical knowledge after which knowledge replacing technical change starts affecting the composition of the set of current core innovators. All else equal when less than twenty percent of the technical knowledge produced in the last 5 years persist on the current main technological trajectory the probability of survival for current fully specialized core innovators start decreasing. This confirms our fourth hypothesis but only with respect of fully specialized core innovators.



**Figure 37: Adjusted predictions of Recent Knowledge Persistence (Main Trajectories)**

#### *Hypothesis 5: Interaction between knowledge persistence, breadth and accumulation*

We have just learned that, when we split the sample into fully specialized and diversified players, persistence keeps being significant for the former and turn insignificant for the latter. This is a first evidence in support of our last hypothesis. A more formal test comes from the estimation of the average marginal effects (AMEs) of diversification at different persistence levels. We estimated the AMEs for increasing levels of persistence ranging from zero to one, with increments of 0.05. We estimated them for each KB bin category with respect to the first bin, which included fully specialized players only and the fifth one, which is the first reliable bin category for intermediate levels of KB. This way we can assess whether increasing KB starting from 0, or 0.5, significantly improves survival chances conditional to the level of knowledge persistence. Table 28 presents the results.

**Table 28: Average Marginal Effects (AMEs) of diversification conditional on the level of knowledge persistence**

| BASE LINE KB BIN 1 (KB=0)        |        |       |       |      |       |       |       |              |
|----------------------------------|--------|-------|-------|------|-------|-------|-------|--------------|
| RK Persistence (MT)              | KB bin | AME   | SE    | z    | p     | CI_lb | CI_ub | Significance |
| 0                                | 7      | 0.361 | 0.151 | 2.39 | 0.017 | 0.065 | 0.657 | **           |
| 0.05                             | 7      | 0.352 | 0.122 | 2.88 | 0.004 | 0.112 | 0.591 | **           |
| 0.1                              | 7      | 0.264 | 0.082 | 3.2  | 0.001 | 0.102 | 0.425 | **           |
| 0.15                             | 7      | 0.157 | 0.064 | 2.47 | 0.014 | 0.032 | 0.281 | **           |
| 0                                | 8      | 0.569 | 0.196 | 2.9  | 0.004 | 0.185 | 0.954 | **           |
| 0.05                             | 8      | 0.484 | 0.147 | 3.3  | 0.001 | 0.196 | 0.771 | **           |
| 0.1                              | 8      | 0.332 | 0.082 | 4.03 | 0     | 0.171 | 0.494 | **           |
| 0.15                             | 8      | 0.188 | 0.068 | 2.77 | 0.006 | 0.055 | 0.321 | **           |
| 0                                | 9      | 0.362 | 0.159 | 2.28 | 0.022 | 0.051 | 0.673 | **           |
| 0.05                             | 9      | 0.353 | 0.134 | 2.64 | 0.008 | 0.091 | 0.614 | **           |
| 0.1                              | 9      | 0.264 | 0.091 | 2.9  | 0.004 | 0.086 | 0.443 | **           |
| 0.15                             | 9      | 0.157 | 0.067 | 2.35 | 0.019 | 0.026 | 0.288 | **           |
| BASE LINE KB BIN 5 (0.4<=KB<0.5) |        |       |       |      |       |       |       |              |
| RK Persistence (MT)              | KB bin | AME   | SE    | z    | p     | CI_lb | CI_ub | Significance |
| 0                                | 7      | 0.380 | 0.169 | 2.25 | 0.025 | 0.048 | 0.711 | **           |
| 0.05                             | 7      | 0.376 | 0.149 | 2.52 | 0.012 | 0.084 | 0.668 | **           |
| 0.1                              | 7      | 0.288 | 0.121 | 2.38 | 0.018 | 0.050 | 0.525 | **           |
| 0                                | 8      | 0.588 | 0.198 | 2.98 | 0.003 | 0.201 | 0.975 | **           |
| 0.05                             | 8      | 0.508 | 0.160 | 3.16 | 0.002 | 0.193 | 0.823 | **           |
| 0.1                              | 8      | 0.356 | 0.117 | 3.05 | 0.002 | 0.127 | 0.585 | **           |
| 0.15                             | 8      | 0.205 | 0.100 | 2.06 | 0.039 | 0.010 | 0.401 | **           |
| 0                                | 9      | 0.381 | 0.173 | 2.2  | 0.028 | 0.042 | 0.720 | **           |
| 0.05                             | 9      | 0.377 | 0.156 | 2.42 | 0.015 | 0.072 | 0.682 | **           |
| 0.1                              | 9      | 0.288 | 0.125 | 2.3  | 0.022 | 0.042 | 0.534 | **           |

To keep the table easily readable we only report AMEs statistically different from zero at the  $p < 0.05$  level. The first half of the table shows that, all else equal, the spread of survival chances compared to fully specialized players (i.e. the AME) is positive and significantly different from zero only for bins 7, 8 and 9<sup>25</sup> and only for levels of knowledge persistence smaller or equal than 0.15. When knowledge persistence increases diversification does not significantly improves survival chances. The same holds when we compute the AMEs using the fifth bin as base line. Filtering players-specific random effects out (i.e. computing the AMEs using Model 8 from Table 24), largely confirms the results. Only the AMEs for a persistent level of 0.15 turns insignificant. Yet, its p-value is very close to 0.1. Our fifth hypothesis was therefore too mild. The argument that the effect of diversification on survival is *stronger* when technological change is knowledge replacing is partially misleading. Our findings show that diversification *only* affects core innovators' resilience when technological change is strongly knowledge-replacing.

#### *Behaviour of the control variables*

Two additional findings emerge from our estimations. They come from the behaviour of the control variables. First, with the seldom and inconsistent exception of research providers and users in some of the models, there are no statistical differences across business models in terms of

<sup>25</sup> Note that AMEs calculation is based on Model 6 from Table 24. KB bins 2 and 10 were omitted from the regressions as they predicted survival perfectly. Consequently, AMEs for these bins could not be computed.

probability to survive as core innovator compared to being an Integrated Device Manufacturer. Given that IDMs are usually financially much bigger than fabless, foundries and suppliers, this also suggests that there are no financial size effects. Second, fully specialized players with less than 10 years of experience as core innovators (i.e. with a presence on the NMPs for up to 2 periods), have significantly less chances of survival than new fully specialized innovators. This suggests that the negative effects of full specialization are particularly stronger in the medium term.

#### *Comparison with probit and complementary log-log models*

As a robustness check, we estimate the same models reported in Tables Table 24, Table 25 and Table 26 using probit and complementary log-log cross-section regressions. Results are reported in the Appendices A.5.4 and A.5.5. Our findings are remarkably robust to the choice of the binary outcome regression model. None of the conclusions derived for our five hypothesis changes using the alternative estimation methods.

## **5.9 Conclusions**

In this work, we analysed the effect of technical knowledge diversification and accumulation on firms' innovative survival in the semiconductor industry. We hypothesized that diversification, accumulation and their joint occurrence are associated with firms' survival in the set of core innovators. Moreover, we argued that knowledge-replacing technical change negatively affects firms' innovative survival and that the diversification and accumulation-induced survival premium should be larger in periods of technological turbulence.

We have proven that indeed diversification, significantly and positively affects innovative survival in the semiconductor industry, though only for large levels of knowledge breadth. In contrast, knowledge accumulation alone does not help to improve the probability to persist as core innovator. We also proved that when technological change is knowledge-replacing (i.e. when recently created technical knowledge is of no use for mastering the current engineering solutions), there is a higher hazard of innovative exit and the set of core innovators undergoes a phase of shakeout. In particular, our findings showed that diversification provides a premium in terms of resilience to changes in the main technological trajectories compared to specialization only when technical knowledge persistence is low. For high levels of persistence, differences in terms of knowledge breadth across the full sample of core innovators do not affect the probability of survival. However, for players that are already mildly diversified, further increasing their knowledge breadth further widening their knowledge breadth provides a survival premium that is not conditional to the level of knowledge persistence. Our findings also showed that these conclusions holds after controlling for the size of the technical knowledge base and for the business model followed by core innovators (IDMs, foundries, fabless, suppliers, users and research providers). This suggests that the benefits of technical knowledge diversification are independent from product diversification and from the stage of value-chain fragmentation. Finally, we also showed that the hazard of exit is particularly strong for mid-levels of innovative experience.

These findings have important managerial and policy implications. Concerning the former, it is clear that even in an industry characterized by large technology modularity, functional or product specialization should not necessarily lead to technical knowledge specialization. When knowledge persistence is low, knowledge specialization increases the hazard of being locked-out from the new technological developments, putting long-run market survival at risk. This raises the question of how to pursuit knowledge diversification when technology is highly modular.

Since in such a technological space it might not be obvious to move across domains, as they share little knowledge and capabilities, it might be necessary to heavily rely on external sources of knowledge, especially for small players. The history of the global semiconductor industry has shown how R&D alliances and technology licensing have been instrumental in the catching-up processes of latecomer firms. Our results suggest that they might also have allowed firms to diversify into poorly related technological domains. This hypothesis can serve as a base for future research.

The policy implications of our findings are particularly important for small catching-up countries characterized by few firms with limited R&D resources. Such firms might face serious difficulties to pursue a strategy of technology diversification and knowledge accumulation at the same time. How can the government help? There are two policy options. First, diversification at the country level can be achieved through specialization at the firm level. When firms' specialization patterns are not overlapping (i.e. when firms are specialized in different areas), country's knowledge is diversified. Such strategy can be called '*diversified specialization*'. Such outcome can hardly emerge through self-organization as Marshallian externalities, coming from knowledge spillovers and economies of co-specialization, might induce firms to focus on similar technological domains. Therefore, diversified specialization requires external coordination by the local government, which should provide incentives for firms to conduct R&D in non-overlapping areas. This would allow achieving, at the country level, Jacobian externalities coming from related variety. As we show in the Appendix A.5.7, the data suggests that this aggregate diversification pattern was achieved in S.Korea and Taiwan. However, further work is necessary to confirm this. A second policy option has more to do with the structure of the technological space. We argued that modularity and knowledge-replacing technical change represent evolutionary traps that expose specialized innovators at risk of innovative extinction. The drop in knowledge persistence urges players to upgrade their knowledge to master emerging domains but high modularity makes knowledge less portable across the technological space. The second policy option in this scenario requires direct government intervention. Public research centres can engage in R&D project that lay at the interface of technological domains. By doing so, the government is providing a technological infrastructure that connects otherwise poorly connected areas in the technology space. The rationale is similar to the argument that sees the provision of physical infrastructure (like railroads or highways) connecting isolated cities or regions as a government's task. Similarly, venturing into research projects whose goal is to create bridges across technology domains is too costly, and possibly not profitable in the short to medium run, to be performed by private firms. Reasons to explore inter-domain connections might not be obvious for profit maximizing firms. Government intervention, in this second policy option, should focus on creating escape paths that can be travelled to move from declining domains to emerging ones.

A major limitation of our study comes from the lack of financial information about the companies in the sample and of data able to better discriminate firms-specific technical capabilities. In particular, data on R&D investments, number of researchers, years of experience of the researchers, being engaged in a strategic alliance, would have greatly helped improve the precision of the estimation of survival chances for fully specialized players, which suffered from a very large variance. In absence of such data, we need to accept that survival of fully specialized players cannot be explained with sufficient confidence just by looking at their level of knowledge accumulation. A second limitation comes from the nature of the study. By focusing only on the semiconductor industry, we are unable to compare our findings with alternative cases characterized by lower knowledge modularity. However, this approach can be easily applied to multi-

industry studies, as long as we can safely rely on their propensity to patent. This provides a second field of future research.

# 6 CONCLUSIONS

In this PhD Thesis, we analysed the complex relationship between technology, industrial dynamics and catching-up in the Semiconductor industry. We did so by developing a new theoretical and methodological framework to study the evolution of the space of engineering solutions along two dimensions, the structure of the system of engineering problems and the paths of problem solving approaches. We argued that changes in these two dimensions are intertwined with the evolution of technological trajectories that are defined at the product level. This novel theoretical framework, and the network analysis methods developed to analyse it empirically, provided the possibility to quantitatively investigate the co-evolution of technology and firms' inventive strategies with previously unmatched resolution and level of statistical rigour. In this conclusive chapter, we summarize the main contributions and findings of the thesis, discuss the implications of this PhD dissertation on evolutionary economics and innovation studies and briefly present its key policy and managerial implications. Finally, we introduce the research agenda sparked by this work.

## 6.1 Summary of the thesis' contributions and findings

This thesis provides a variety of methodological contributions to the analysis of technology dynamics and catching-up. We designed a set of patent citation network analysis methods to unfold sources and speed of technical improvements of catching-up firms (Ch.2), measure the stability of the ranking of engineering problems and the paths of problem-solving approaches (Ch.3), identify firms that are able to influence the direction of technological change (Ch.3), reveal significant path- and ranking -changers and followers (Ch.3), identify technology domains, trace their life cycle and measure firms dynamic technical comparative advantages (Ch.4), measure the extent to which technical change is knowledge replacing (Ch.5) measure firms' technical knowledge breadth and depth (Ch.5), and quantify the premium in innovative survival chances associated with firms' technical knowledge diversification and accumulation (Ch.5). We applied these methodologies to study technology and industrial dynamics in the semiconductors using patent data from the USPTO-NBER database between 1976 and 2006. The choice of the Semiconductor industry as field of study allowed validating the findings of our methodologies against the breadth of literature on latecomer catching-up strategies that builds on anecdotal

evidences, policy analysis and case studies. Our quantitative methods proved to be able to replicate these qualitative evidences.

New theoretical insights on the co-evolution of technology and firms' inventive strategies in the Semiconductors were achieved. The network analysis of semiconductor patent citations also shed light on the paths followed by firms from successful latecomer countries (Taiwan, Korea and Singapore). In Chapter 2, we unfold how, despite sharing relatively similar initial national conditions in terms of technology environment at the time of their first USPTO patents, firms from successful and unsuccessful catching-up countries followed different inventive strategies with respect of the sources and speed of technical improvements. In Taiwan, Singapore and Korea, we could observe the rise of an industry, with many players that progressively learned from each other and adopted a strategy of rapid technical improvements. In Chapter 3, we showed that the main paths of engineering improvements have been disrupted twice in the last four decades. A temporary but substantial disruption in the approaches to problems solving occurred in the first half of the 1990s. This change was caused by attempts to look for alternative ways of improving integrating circuits' computational power and energy efficiency. One of such attempts focused on circumventing second order problems caused by miniaturization, like power leakage due to heat dispersion. A second one explored the feasibility of stacking dies vertically rather than horizontally, to achieve a higher density of transistors on a chip. A milder change in the direction of the main paths of engineering improvements was observed in the beginning of the 2000s. This discontinuity was coupled to a change in the ranking of engineering problems, caused by the diffusion of new applications for semiconductor devices, such as LCD TV and computer monitors, tablets and smartphones. Moreover, further improvements in metal-oxide-semiconductor field-effect transistors (MOSFETs) also stand behind these changes. Our findings also show that, in the 1990s and first half of the 2000s, firms from successful latecomer countries, especially Taiwanese and Singaporean foundries, were significantly more path and ranking followers than expected by their size and time of entry. Some of them, notably those engaged in LCD technology, proved to be significant ranking changers in the early 2000s. In most of the cases, their strategy to initially follow well-established paths of improvements and respect the prevailing perception on the ranking of engineering problems has granted them a very influential role. Their patents were significantly more central in the network of technical improvements than expected by firm size and time of entrance. The analysis in Chapter 4 confirmed that latecomer countries were indeed able to upgrade their technical knowledge during the process of catching-up. Their technical comparative advantage initially laid in improving relatively mature and declining technologies. However, in the course of about a decade, it evolved into emerging technology domains, most notably in breakthroughs related to LCD technology. Our method to identify the life cycle stage of technology domains, developed in Chapter 4, also allowed investigating significant differences in the comparative advantage of new innovators and incumbents along the life cycle. In accordance with what hypothesized in the literature on disruptive technology (e.g. Christensen, 1997) and industry life-cycles (e.g. Klepper, 1997), our findings revealed that new innovators have a slightly higher comparative advantage than incumbent in breakthrough technology domains and a much larger one in disruptive technologies. Finally, in Chapter 5, we showed that technical knowledge is highly modular in the Semiconductor industry. This makes knowledge and experience technology-specifics and therefore less portable across technology domains. Due to this configuration of the system, technical knowledge diversification provides better chances of innovative survival than specialization, in particular when technological change is knowledge-replacing. We also found that the knowledge base of semiconductor foundries from Taiwan and Singapore is broadly diversified, perhaps due to their role



of generalist manufacturers of microchips for a large set of chip-designers. This allowed them persisting in the club of core innovators uninterruptedly from the mid-1990s until at least 2006.

## 6.2 Implications of this PhD dissertation for technology and innovation studies

What do the findings of this thesis collectively tell us about technology development dynamics, inventive strategies and catching-up pathways? A coherent story emerges from our empirical analysis, which matches anecdotal evidences presented in the literature (Brown and Linden 2009, Bell and Juma, 2008, Chang and Tsai, 2002, Lee and Lim 2001, Hobday 2000, Langlois and Steinmueller 1999, Mathews and Cho 1999, Cho et al., 1998, Chang et al., 1994). Catching-up firms from successful latecomer countries, initially preferentially focused on rapidly improving foreign and relatively obsolete technologies characterized by long and established engineering design trajectories. They then progressively started learning from national sources, fostering knowledge diffusion and industry building within their country, and upgrading their knowledge by shifting their comparative technical advantage to younger and emerging technology domains. At the same time, firms from these countries built a broadly diversified knowledge base, which allowed them to spread development risks and secure their role as core innovators. An obvious question emerges. *Is this catching-up trajectory, which emerged from the careful analysis of patent records, the only possible one that could have allowed rapid development of frontier technical capabilities?* As we discussed in the introduction of this thesis, attempting to answer this question is not only hard but it is likely to provide misleading messages. Because of the complex nature of technical and economic change, universally valid and replicable strategies hardly exist. However, the description of how the technological environment of the Semiconductor industry evolved and the analysis of what made successful latecomer firms special, allow better assessing if similar scenarios may exist in other industries in different moment in times. Therefore, we do not have a recipe for catching-up, but we at least have a carefully described benchmark. Such benchmark has a number of important implications for innovation studies and evolutionary economic theory.

Firstly, we showed how the interaction between demand, product evolution and changes in the space of engineering problems can explain the rise and decline of technology domains. We described technology domains as areas of research that define a set of common technological problems that are tackled applying similar mindsets and toolboxes. As such, technology domains can span product and component boundaries. In other words, the evolution of different products may depend on the solution of similar engineering challenges. We have discussed how this is the case in the Semiconductor industry, where, due to the ubiquitous application of semiconductors in modern electronics, solving some key design challenges related to their miniaturization affect performance improvements of several products. The key question then becomes, *how do technology domains evolve?* In this thesis, we answered it by discussing how two variables describe the entire life cycle of technology domains. These are the amount of firms' inventive effort devoted to solving the central engineering problems behind a domain and the level of firms' exploration of unconventional problem solving approaches. These two variables can describe whether a technology domain is emerging, maturing, declining or renewing, as we showed in Chapter 4. The theoretical description of the life cycle of technology domains and the development of methodologies to empirically analyse it are a major contribution of this work. In fact, as we discussed, because technology domains affect different products one can now envisage ways of explaining and, possibly, predicting the life cycle of products and industries by looking at their en-

gineering design trajectories. These future research agenda is discussed in more details in the last section.

Secondly, by linking evolutionary dynamics of products and technology domains, the theoretical framework developed in this dissertation can potentially explain the clustering of big innovations in time, as theorize by Schumpeter and reported by Silverberg and Verspagen (2005). Therefore, it can contribute to identify a key driver of economic cycles. Realize that, at its very essence, technological progress is achieved by searching for the right approach to solve engineering challenges allows us to focus on the cognitive process behind the search of design solutions. Moreover, acknowledging that the space of engineering problems has a structure is the first step toward analysing how solutions propagate. Eventually, this can explain why we observe waves of new products. The analysis of the semiconductor industry, which we have developed in this thesis, shows how there is a cyclic dependence between product evolution, engineering design trajectories and the evolution of technology domains. For instance, we have discussed how semiconductor devices (such as integrated circuits and memories) have evolved around a well-defined technological trajectory. Miniaturization is the product feature that must be imperatively achieved by any player in the industry. Miniaturizing microchips simultaneously allows increasing computing power and storage space. Therefore, the industry has certainly evolved along a technological trajectory that clearly defined the features of semiconductor devices that need to be improved. Determining whether miniaturization is the result of consumer preferences or designers' individual choices was beyond the scope of this work. Our goal was simply to analyse how the miniaturization technological trajectory has determined which design challenges needed to be addressed and how solutions were searched. Making microchips smaller and smaller led to the emergence of a particular set of engineering problems, such as heat dispersion and deterioration in microchips. We showed how the search for solutions to these problems mostly followed established problem solving approaches except for a brief period in the mid-1990s coinciding with the entrance of latecomers. However, the miniaturization also allowed new products to emerge. These products introduced new design challenges. One example of this is the application of light-emitting diodes (LCD) to displays. Ultimately new and better products (like flat screen televisions and monitors or small portable electronic devices) emerged because of miniaturization advances made possible by solving design challenges. We showed, in Chapter 3 and 4, how the structure of engineering problems and their ranking of importance evolved during this process because some players started devoting a larger share of inventive effort to searching solutions to these new challenges. These dynamics clearly show how product features, engineering design challenges and firms' search for better approaches to tackle engineering bottlenecks, co-evolve in a complex manner. It is precisely the process of searching across problem solving approaches that can lead to solutions that are able to expand the space of what is technically feasible. In particular, when solutions to engineering problems that are ubiquitous across products are found, like those affecting the miniaturization of microchips, waves of new products can emerge.

Thirdly, the evolution of engineering design trajectories and technology domains can explain the emergence of catching-up pathways in an industry. Firms can apply different inventive strategies, which can be classified based on which engineering problems they try to tackle and in which way. Usually industries are characterized in the literature as relatively homogeneous entities that can be described by a single life cycle. This compresses information on the variety of technology domains that is present in a given industry. We showed that technology domains that are at different stages of their life-cycle co-exist in the Semiconductor industry. This implies that firms can choose among a variety of strategies, which may lead to different catching-up out-

comes. The strategy mix can be built around two dimensions. First, a firm has to decide which engineering challenges to tackle, i.e. in which technology domains she wants to focus on. Second, the firm needs to decide which explorative strategy to follow to navigate across the space of problem solving approaches. Expected payoff and their variance may differ according to the combination of strategies. For instance, one could focus on emerging domains while following a comparatively more conservative exploration strategy, or devote inventive effort to mature areas while pursuing new approaches to problem solving. We showed how successful catching-up firms evolved from a very risk-adverse strategy focused on relatively old technology domains and path-following problem solving to a riskier but potentially more rewarding comparative advantage in breakthrough technology domains, while still maintaining a relatively preference for conservative problem-solving approaches. However, this does not preclude the possibility that other strategies would have been effective. A more systematic study of the expected payoff and variance for each combination of strategies across industries is required to provide a sound and credible answer to this question. However, regardless of whether other strategies could have been equally viable and profitable, it is interesting to discuss why latecomers opted, at least initially, for a strategy mix based on devoting inventive effort to the then prevailing engineering challenges while following established problem-solving approaches and specializing in mature and renewing technology domains, as shown in Chapters 3 and 4. The common explanation that emerges from the literature (e.g. Lee and Lim, 2001) interprets following the path as a signal of progressive learning. However, there may be an alternative explanation, which emerged from an interview that the author has recently conducted to a former high-level technology development manager at Chartered Semiconductor Manufacturing, Professor Lap Chan, from Singapore University of Technology and Design, who has three decades of working experience in the Semiconductor industry, two of which in Singapore. A lot of the senior and junior managers in the newly established semiconductor firms in Singapore and Taiwan were people that had been trained in the US, as also confirmed by Chang and colleagues (1994) as well as Bell and Juma (2008). Many of them also worked for a significant period for top US companies. Therefore, they did not really need to learn from scratch. Initially local firms were not engaged in heavy patenting. However, as they started producing devices outsourced by US firms and using foreign developed manufacturing processes, they also started receiving the first requests of paying high licensing fees. This made them realize that, in order to survive, they needed to pile up patents as defensive weapons for cross-licensing agreements. This became a common strategy in the global Semiconductor industry, as discussed by Hall and Ziedonis (2001). To do that, they had to choose in which technology domains they needed to focus their inventive effort. The natural choice was to select the engineering challenges that were then prevailing in the industry, namely relatively mature domains, as they provided better defensive value. Because many of the employees involved in the R&D process had been trained and worked in the US, the most obvious way of tackling such challenges was to apply the problem solving approaches that they were exposed to while in the States. This would explain why these firms were significant path-followers, focused on mature and renewing domains, as shown in Chapters 3 and 4. We can then speculate that the change in national and development context may have initially disconnected these inventors with the direction of technical knowledge evolution in the industry. Furthermore, emerging domains may have been too risky, for newly founded firms in latecomer countries, that needed to quickly pile up patents for cross-licensing. This, could explain why these companies shifted their focus to emerging domains only later, at the beginning of the 2000s. Incidentally, it may also shed light on the apparently puzzling fact that, despite being specialized in manufacturing, foundries from Taiwan and Singapore are highly diversified across technology domains, including in areas that

are design-intensive, as shown in Chapter 5. In fact, if looked from this perspective, latecomer firms' wide technical knowledge diversification may take a different meaning. In Chapter 5 we showed how early technical knowledge diversification is crucial for survival at the technological frontier. We provided three interpretations of this finding in the dissertation. First, knowledge diversification provides inventive risk sharing. Even for functionally specialized or mono-product firms, being diversified at the technical knowledge level allows spreading the risk of knowledge obsolescence across domains. Second, it facilitates exploring in which domains the firm has a comparative advantage. Third, it provides absorptive capacity. Knowledge diversification for firms specialized in manufacturing makes it easier to interact with many design partners. However, as it emerged from the discussion with industry experts, there could be a fourth alternative explanation. Early technical knowledge diversification allowed latecomer companies in the Semiconductors to build up wide cross-licensing portfolio for defensive purposes. This strengthened their bargaining position when dealing with design partners operating in different technology domains. This last hypothesis is much less related to *learning* and has much more to do with *incentives* and *competition*. Further work is necessary to shed light on the theoretical implications of this counterintuitive finding.

### 6.3 Policy and managerial implications

This thesis' findings, theories and methods have a number of practical policy and managerial implications. Understanding the structure of the system of engineering problems and describing the direction of evolution of engineering trajectories is of the outmost importance to design inventive strategies for firms, regions and countries that better fit with the evolution of the technology space. In particular, this dissertation has two types of policy and managerial implications. Firstly, we can now precisely quantify the impact of firms' inventive strategies and of the innovation output of geographic entities' such as cities, regions and countries on future technology developments. Secondly, our analysis allows envisaging a range of strategic options that firms and geographic entities can choose to modify their location with respect to the prevailing engineering trajectories and their comparative advantage pattern across technology domains.

It is of great interest to firms and geographic entities to know their location on the main engineering trajectories and comparing their comparative advantage pattern with competitors. Our network analysis methods are able to measure firms', cities', regions' and countries' inventions' centrality on the main engineering design trajectories, assess the variety of problem-solving approaches used within these entities and quantify their technological comparative advantage across technology domains that are in different life cycle stages. These are direct measures of how firms and geographic entities are performing in terms of technical capability building. They can also be used to identify competitors' inventive strategies. It is worth mentioning that our study does not make any conclusion about which strategy ought to be followed. In particular, we have shown that being path-follower, rather than path-changer, or vice versa, does not predict the amount of further inventions that will connect to a given patent or to patents from a given entity above random expectations (as shown in Figure 45 in the Appendix A.3.2). This means that the future success of a given engineering trajectory does not depend on the centrality of the previous solutions.

However, even though being a path-follower and being specialized in mature and declining domains can be a viable strategy to exploit comparative advantage in the short run, it is reasonable to state that it may endanger the ability of persist innovating at the frontier in the long run. Indeed, in Chapter 5, we have shown how knowledge diversification across domains provides

better survival chances as core innovator. Therefore, it becomes crucial to understand how innovators (be they firms or geographic entities) can turn to path-changers and upgrade their advantage to emerging technology domains. The first task is to identify emerging technology domains and changes in engineering design trajectories. As we showed in this work, this can be done by looking at changes in the main paths of citations in a patent network and on clustering the invention space by similarity in approaches to problem-solving. Once emerging domains and trajectories have been detected, innovators then need to locate themselves in the space of engineering solutions and find the direction to follow to build strength in the desired domain(s). The available options differ, depending on the unit of analysis. Suppose that a path-following firm wants to change the way its engineers tackle design problems and/or the very set of design challenges they focus on. Because learning is path-dependent it can be very difficult to make such turn by relying only on in-house researchers' inventive effort. It may be more viable to hire external talents that have been exposed to different approaches to problem solving and/or working in different technology domains. If attracting such talents is unfeasible, reorganizing internal research teams may be more effective as a way to stimulate novel thinking, as shown by Uzzi et al. (2013). Access to a different pool of problem solving approaches may also be provided by external partnerships of different formal strength, ranging from alliances to mergers and acquisitions. Being measured at the firm level, our path- and rank-changing indices, as well as the analysis of comparative advantage across technology domains, can provide managerial guidance in the process of identifying useful external sources of knowledge for partnerships, acquisitions and even hiring strategies.

Geographic entities have a similar set of strategic options to upgrade their comparative advantage to emerging domains and, if desired, to reposition themselves on more novel engineering design trajectories. External talents could be attracted with appropriate incentives or local talents can be sent for training purposes to other geographic areas that have expertise in emerging technology domains or new problem-solving approaches. Similarly, policies could be designed to foster cross-boundary partnership among firms. The key goal of these policies should be to increase the variety of problem-solving approaches used within the given geographic area and make sure that the comparative advantage is not exclusively focused on mature and declining technology domains. Obviously, this is a much more difficult task for entities that have limited resources. In Chapter 5, we suggested how bridge formation and diversified specialization might facilitate achieving these goals. The former refers to the local or national government action of investing in research projects or laboratories whose goal is the generation of knowledge at the crossroad between technology domains. This would create a path that local firms can follow to move their specialization from one domain to a previously unrelated one. Diversified specialization is a different way of achieving risk sharing at the local or national level while keep having strongly specialized firms. As we discussed in Chapter 5 (see Appendix A.5.7), the data suggests that this was done in S.Korea and Taiwan, where, at least between 1990 and 2006, knowledge diversification at the national level was achieved through knowledge specialization at the firm level with narrow inter-firm knowledge overlap.

In the knowledge valorisation addendum reported at the end of the manuscript, we thoroughly discuss possible consulting service packages addressing policy makers' and practitioners' need to evaluate the impact of their inventive strategies and identify diversification pathways to emerging domains.

## 6.4 Research agenda

A very promising research agenda can be built starting from the theoretical framework developed in this thesis to analyse the co-evolution of technology and firms' inventive strategies and the network analysis methods designed to empirically study this co-evolution. Such agenda is composed of three parts: *causes*, *effects* and *prediction*. First, in the thesis we have shown how to identify changes in engineering trajectories, ranking of design problems and trace the evolution of technology domains. It must be then understood which are the *causes* of these changes. Second, it is equally important to study the *effects* that such phenomena have on industrial dynamics and economic cycles. Third, if a sound understanding of these phenomena is achieved, it may be possible to *predict* upcoming changes in trajectory, the emergence of new technology domains and the decline of old ones.

### 6.4.1 Causes

The indices that we developed in Chapter 3, allow measuring changes in the ranking of engineering problems and the paths of problem-solving approaches, and identifying path- and ranking-changers. Once we identified who path- and ranking-changers are, we can then look for predictors of their strategic behaviour. In other words, we can investigate if firms falling within these categories have a common profile that can explain their strategic choices. We have already shown how firm size, measured by number of patents, is not a good predictor of their inventive strategies (as reported in Figure 45 in the Appendix A.3.2). One possible candidate is the firm's age and its relative performance compared to competitors. Spinoffs and firms that are less successful might have more incentives to explore new paths. An alternative explanation of path-changing and rank-changing behaviour may be provided by the previous working experience and social network of a firm's research workforce. Learning is path-dependent and tacit knowledge is sticky. Therefore, it is reasonable to hypothesize that the way firms tackle engineering challenges depend on what type of problem-solving approaches their researchers have been exposed to during their career, either personally or by social influence from their close peers. The availability of harmonized inventors and firms identification numbers (Li et al., 2014) now makes it possible to pursue this line of research.

### 6.4.2 Effects

Changes in the allocation of inventive effort across engineering problems and in the way firms tackle them can have wider consequences on industrial and economic dynamics. Evolutionary economic theory predicts that large waves of firm entry can be observed when there are changes in the technological regime of an industry (Klepper, 1997 and 1996; Malerba and Orsenigo, 1997). Similarly, one possible line of research could seek to assess whether technology domains that witness a change in their engineering design trajectory also attract more innovators. A further study could focus on investigating how the stability of the ranking of innovators in a given domain is affected by changes in the set of engineering challenges and in the problem solving approaches followed to tackle them.

We have also highlighted how the structure, topology and evolution of the system of technology domains, are important determinants of firms' innovative survival. Furthermore, we showed how new entrants, in general, tend to have a comparative advantage in emerging technology domains, but latecomers tend to initially specialize in following established paths in rela-

tive more mature technologies. This suggests the importance of analysing how technological change propagates. The availability of patent data, coupled with advances in network analysis technique, text mining, latent semantic analysis and computational power, allow envisaging ways to disentangle the system of engineering problems and study change propagation across technology domains. This can potentially explain clustering of big innovations in time and the formation of economic cycles.

Finally, the ability to identify firms' inventive strategies opens the door to the possibility of studying how the expected payoffs are related to the external context. For instance, the technique described in Chapter 4, to identify the life cycle stage of technology domains, can be used to investigate whether the type of diversification strategies that provides better chances to be successful is affected by the stage of development of a technology domain, or how life cycles of related domains affect each other's fate.

### 6.4.3 Predictions

A common criticism that economists hear from the public is the poor predicting performance of many economic theories, especially when abrupt economic change occur. Therefore, an obvious line of research that could follow up this thesis work addresses the possibility of developing a predicting model of changes in the structure and ranking of the system of engineering problems and in the engineering design trajectories followed in an industry. The first two parts of the research agenda focused on describing studies that can potentially improve our understanding of what causes changes in engineering trajectory and in the space of technology domains and which effect these changes have on industrial and economic dynamics. Reaching a sound understanding of causes and effects of technical change would allow us to design studies that would seek to identify early signals of changes in engineering trajectories and in the relative importance of engineering problems and predict their effects. For instance, if one has identified which factors affect firms' decision to allocate inventive effort to new design challenges or to start exploring alternative paths to search for design solutions, it becomes possible to estimate the probability that these phenomena will happen. In addition, knowledge of the effects that these phenomena have on industrial and economic dynamics can help to associate probabilities of occurrence for other events, such as industry shakeouts, or to predict rates of performance improvements for different technologies, given the characteristics of their engineering design trajectory.

The technical feasibility of pursuing this rich and ambitious research agenda makes us optimistic about the existence of something that eminent scholars described as the *PhD afterlife*.





# 7 REFERENCES

- Abernathy, W. J., & Utterback, J. M. (1978). Patterns of industrial innovation. *Technology review*, 64, 254-228.
- Abramovitz M. (1994), "The Origins of the Postwar Catch-Up and Convergence Boom", in Fagerberg J., Verspagen B. and von Tunzelman N. (eds), "The Dynamics of Technology, Trade and Growth", Aldershot, UK, Edward Elgar, pp. 21-52.
- Adams, P., Fontana, R., Malerba F. (2013), The Magnitude of Innovation by Demand in a Sectoral System: The Role of Industrial Users in Semiconductors. *Research Policy*, 42, p.1-14.
- Agarwal, R., & Gort, M. (2002). Firm and product life cycles and firm survival. *American Economic Review*, 184-190.
- Alcacer, J., & Gittelman, M. (2006). Patent citations as a measure of knowledge flows: The influence of examiner citations. *The Review of Economics and Statistics*, 88(4), 774-779.
- Alstott, J., Triulzi, G., Yan, B., Luo, J. (2015), Mapping Technology Space by Normalizing Technology Networks. arXiv:1509.07285
- Antonelli, C. (2003). The economics of innovation, new technologies and structural change. Routledge.
- Antonelli, C., Crespi, F., & Scellato, G. (2012). Inside innovation persistence: New evidence from Italian micro-data. *Structural Change and Economic Dynamics*, 23(4), 341-353.
- Arthur, W. B. (2009). The nature of technology: What it is and how it evolves. Simon and Schuster.
- Audretsch, D. B. (1995). Innovation, growth and survival. *International journal of industrial organization*, 13(4), 441-457.
- Balassa, B. (1965), Trade Liberalization and 'Revealed' Comparative Advantage', *The Manchester School of Economic and Social Studies*, Vol. 32, pp. 99-123.
- Baldwin, C. Y., & Clark, K. B. (2000). *Design rules: The power of modularity* (Vol. 1). MIT press.
- Baofu, P. (2007). The future of complexity: conceiving a better way to understand order and chaos. World Scientific.
- Batagelj V. (2003), "Efficient Algorithms for Citation Network Analysis", University of Ljubljana, Department of Mathematics, cs.DL/0309023.
- Bayus, B. L., & Agarwal, R. (2007). The role of pre-entry experience, entry timing, and product technology strategies in explaining firm survival. *Management Science*, 53(12), 1887-1902.
- Bekkers R., Martinelli A. (2010), "The interplay between standardization and technological change: A study on wireless technologies, technological trajectories, and essential patent claims", Paper presented at the Druid Summer conference 2010.

- Bell Jr, B. W., & Juma, C. (2008). Institutional reform and technology development: the case of ITRI. *International Journal of Technology and Globalisation*, 4(3), 296-313.
- Bergek, A., Berggren, C., Magnusson, T., & Hobday, M. (2013). Technological discontinuities and the challenge for incumbent firms: Destruction, disruption or creative accumulation?. *Research Policy*, 42(6-7), 1210-1224.
- Boschma, R., & Frenken, K. (2011). The emerging empirics of evolutionary economic geography. *Journal of Economic Geography*, 11(2), 295-307.
- Borota, T. (2012). Innovation and imitation in a model of North-South trade. *Journal of International Economics*, 87(2), 365-376.
- Breschi, S., Lissoni, F., & Malerba, F. (2003). Knowledge-relatedness in firm technological diversification. *Research Policy*, 32(1), 69-87.
- Breschi, S., Malerba, F., & Orsenigo, L. (2000). Technological regimes and Schumpeterian patterns of innovation. *The Economic Journal*, 110(463), 388-410.
- Brown, C., & Linden, G. (2009). Chips and change: how crisis reshapes the semiconductor industry. MIT Press, Cambridge, Massachusetts.
- Brusoni, S., Criscuolo, P., & Geuna, A. (2005). The knowledge bases of the world's largest pharmaceutical groups: what do patent citations to non-patent literature reveal?. *Economics of Innovation and New Technology*, 14(5), 395-415.
- Brusoni, S., & Geuna, A. (2003). An international comparison of sectoral knowledge bases: persistence and integration in the pharmaceutical industry. *Research Policy*, 32(10), 1897-1912.
- Brusoni, S., Prencipe, A., & Pavitt, K. (2001). Knowledge specialization, organizational coupling, and the boundaries of the firm: why do firms know more than they make?. *Administrative science quarterly*, 46(4), 597-621.
- Buenstorf, G. (2007). Evolution on the shoulders of giants: entrepreneurship and firm survival in the German laser industry. *Review of Industrial Organization*, 30(3), 179-202.
- Camerani R., & Malerba, F. (2007). Patterns of Technological Entry in Different Fields: an Analysis of Patent Data. Report to the European Commission. FP6 Project: Knowledge-Based Entrepreneurship: Innovation, Networks and Systems.
- Cefis, E., & Marsili, O. (2005). A matter of life and death: innovation and firm survival. *Industrial and corporate change*, 14(6), 1167-1192.
- Chang, P. L., Shih, C., & Hsu, C. W. (1994). The formation process of Taiwan's IC industry—method of technology transfer. *Technovation*, 14(3), 161-171.
- Chang, P. L., & Tsai, C. T. (2002). Finding the niche position—competition strategy of Taiwan's IC design industry. *Technovation*, 22(2), 101-111.
- Chang, S. C. (2005). The TFT-LCD industry in Taiwan: Competitive advantages and future developments. *Technology in Society*, 27(2), 199-215.
- Cho, D. S., Kim, D. J., & Rhee, D. K. (1998). Latecomer strategies: evidence from the semiconductor industry in Japan and Korea. *Organization science*, 9(4), 489-505.
- Christensen, C. M. (1993). The rigid disk drive industry: a history of commercial and technological turbulence. *Business history review*, 67(4), 531-588.
- Christensen, C. (1997). The innovator's dilemma: when new technologies cause great firms to fail. Harvard Business School Press.
- Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative science quarterly*, 128-152.
- Cottrell, T., & Nault, B. R. (2004). Product variety and firm survival in the microcomputer software industry. *Strategic Management Journal*, 25(10), 1005-1025.
- Criscuolo, P., & Verspagen, B. (2008). Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Research Policy*, 37(10), 1892-1908.
- D'Este, P. (2005). How do firms' knowledge bases affect intra-industry heterogeneity?: An analysis of the Spanish pharmaceutical industry. *Research Policy*, 34(1), 33-45.

- Dosi G. (1982), "Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change", *Research Policy*, vol. 11, no. 3, June, pp. 147-62.
- De Weck, O. L., Roos, D., & Magee, C. L. (2011). Engineering systems: meeting human needs in a complex technological world. MIT Press.
- Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education*, 94(1), 103-120.
- Ederington, J., & McCalman, P. (2011). Infant industry protection and industrial dynamics. *Journal of International Economics*, 84(1), 37-47.
- Epicoco, M. (2012). Knowledge patterns and sources of leadership: Mapping the semiconductor miniaturization trajectory. *Research Policy*.
- Ernst D. (2005), "Internationalization of Innovation: Why is Chip Design Moving to Asia?", *International Journal of Innovation Management*, Vol. 9, No. 1, pp. 47-73.
- Ernst, D. (2005a). Complexity and internationalisation of innovation—why is chip design moving to Asia?. *International Journal of Innovation Management*, 9(01), 47-73.
- Ernst, D. (2005b). Limits to modularity: reflections on recent developments in chip design. *Industry and Innovation*, 12(3), 303-335.
- Fagerberg J. and Godinho M. (2005), "Innovation and Catching-up", in Fagerberg, J., Mowery, D., and Nelson, R., (eds.), "Oxford Handbook of Innovation", Oxford Un Press, Oxford, 2004.
- Fleming, L. (2012). Breakthroughs and the "long tail" of innovation. *MIT Sloan Management Review*, 49(1).
- Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: evidence from patent data. *Research policy*, 30(7), 1019-1039.
- Fontana, M. (2014). Pluralism (s) in economics: lessons from complexity and innovation. A review paper. *Journal of Evolutionary Economics*, 24(1), 189-204.
- Fontana, M. (2010). The Santa Fe Perspective on economics: Emerging patterns in the science of complexity. *History of Economic Ideas*, 18(2), 167.
- Fontana, R., & Nesta, L. (2009). Product innovation and survival in a high-tech industry. *Review of Industrial Organization*, 34(4), 287-306.
- Fontana R., Nuvolari A., Verspagen B. (2009), "Mapping technological trajectories as patent citation networks. An application to data communication standards", *Economics of Innovation and New Technology*, Vol. 18 (4), pp. 311-336.
- Fontana, R., & Malerba, F. (2010). Demand as a source of entry and the survival of new semiconductor firms. *Industrial and Corporate Change*, 19(5), 1629-1654.
- Foster, J. (2005). From simplistic to complex systems in economics. *Cambridge Journal of Economics*, 29(6), 873-892.
- Freeman, C., & Soete, L. (1997). The economics of industrial innovation. Psychology Press.
- Frenken, K. (2006). Technological innovation and complexity theory. *Economics of Innovation and New Technology*, 15(2), 137-155.
- Frenken, K., & Boschma, R. A. (2007). A theoretical framework for evolutionary economic geography: industrial dynamics and urban growth as a branching process. *Journal of Economic Geography*.
- Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional studies*, 41(5), 685-697.
- Gambardella, A., Harhoff, D., & Verspagen, B. (2008). The value of European patents. *European Management Review*, 5(2), 69-84.
- Garfield E. (1979), "Citation Index – Its theory and application in Science, Technology and Humanities", Philadelphia: Institute for Scientific Information Press.
- Garfield E., Sher I., Torpie R.J. (1964), "The Use of Citation Data in Writing the History of Science", Philadelphia: Institute for Scientific Information Press.

- Gerschenkron A. (1962), "Economic Backwardness in Historical Perspective", The Belknap Press, Cambridge MA.
- Giffin, M., de Weck, O., Bounova, G., Keller, R., Eckert, C., & Clarkson, P. J. (2009). Change propagation analysis in complex technical systems. *Journal of Mechanical Design*, 131(8), 081001.
- Goode R. (1959), "Adding to the Stock of Physical and Human Capital", *The American Economic Review*, Vol. 49(2), pp. 147-155.
- Grossman G.M. and Helpman E. (1991), "Quality Ladders and Product Cycles", *The Quarterly Journal of Economics*, Vol. 106, pp.
- Grossman, G. M., & Helpman, E. (2005). Outsourcing in a global economy. *The Review of Economic Studies*, 72(1), 135-159.
- Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools", NBER Working Paper 8498.
- Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of economics*, 16-38.
- Hall B.H. and Ziedonis R.H. (2001), "The patent paradox revisited: an empirical study of patenting in the US semiconductor industry, 1979-1995", *RAND Journal of Economics*, Vol. 32, No. 1, pp. 101-128.
- Henderson, R. M., & Clark, K. B. (1990). Architectural innovation: the reconfiguration of existing product technologies and the failure of established firms. *Administrative science quarterly*, 9-30.
- Hobday M. (2000), "East versus Southeast Asian innovation systems: Comparing OEM- and TNC- led growth in electronics", in L. Kim and R. Nelson (Eds.), "Technology, learning, and innovation: Experiences of newly industrializing economies", Cambridge & New York: Cambridge University Press.
- Hummon N.P., Doreian P. (1989). Connectivity in a citation network: The development of DNA theory. *Social Networks*, 11, 39-63.
- Hung, S. W. (2006). Competitive strategies for Taiwan's thin film transistor-liquid crystal display (TFT-LCD) industry. *Technology in Society*, 28(3), 349-361.
- Jaffe, A. B., & Trajtenberg, M. (2002). Patents, citations, and innovations: A window on the knowledge economy. MIT press.
- Jovanovic, B., & Rousseau, P. L. (2005). General purpose technologies. *Handbook of economic growth*, 1, 1181-1224.
- Karniouchina, E. V., Carson, S. J., Short, J. C., & Ketchen, D. J. (2013). Extending the firm vs. industry debate: Does industry life cycle stage matter?. *Strategic Management Journal*, 34(8), 1010-1018.
- Kauffman, S. A. (1988). The evolution of economic webs. Philip W. Anderson, Kenneth J. Arrow, and David Pines, eds., *The economy as an evolving complex system*, Addison-Wesley, Reading, Mass, 125-146.
- Kuhn T. (1959), "The Structure of Scientific Revolution", Chicago: Chicago University Press.
- Kim, L., & Nelson, R. R. (Eds.). (2000). Technology, learning, and innovation: Experiences of newly industrializing economies. Cambridge University Press.
- Klepper, S. (1996). Entry, exit, growth, and innovation over the product life cycle. *The American economic review*, Vol. 86, No. 3, p. 562-583.
- Klepper, S. (1997). Industry life cycles. *Industrial and corporate change*, Vol. 6, No. 1, pp. 145-182.
- Klepper, S. (2002). The capabilities of new firms and the evolution of the US automobile industry. *Industrial and corporate change*, 11(4), 645-666.
- Klimek, P., Hausmann, R., & Thurner, S. (2012). Empirical confirmation of creative destruction from world trade data. *PloS one*, 7(6).
- Koh, H., & Magee, C. L. (2006). A functional approach for studying technological progress: Application to information technology. *Technological Forecasting and Social Change*, 73(9), 1061-1083.
- Krafft, J., & Quatraro, F. (2011). The dynamics of technological knowledge: from linearity to recombination. 2011) *Handbook on System Dynamics of Technological Change*. Cheltenham: Edward Elgar, 181-200.

- Kristensen T. (1974), "Development in Rich and Poor Countries", New York: Praeger.
- ICE (Integrated Circuit Engineering Corporation) (1996), Status report 1996, National Museum of American History's Smithsonian Chip Collection, retrieved online on 19/05/2013 at: <http://smithsonianchips.si.edu/ice/cd/CEICM/SECTION1.pdf>
- Langlois, R. N., & Steinmueller, W. E. (1999). The evolution of competitive advantage in the worldwide semiconductor industry, 1947-1996. *The Sources of Industrial Leadership*. Cambridge University Press: New York, 19-78.
- Lall, S. (1993). Understanding technology development. *Development and change*, 24(4), 719-753.
- Lee, K. (2014). Schumpeterian analysis of Economic Catch-up. Knowledge, Path-Creation, and Middle-Income Trap. Cambridge University Press.
- Lee, K., & Lim, C. (2001). Technological regimes, catching-up and leapfrogging: findings from the Korean industries. *Research policy*, 30(3), 459-483.
- Lee, K., Lim, C., & Song, W. (2005). Emerging digital technology as a window of opportunity and technological leapfrogging: catch-up in digital TV by the Korean firms. *International Journal of Technology Management*, 29(1), 40-63.
- Levinthal, D. A., & March, J. G. (1993). The myopia of learning. *Strategic management journal*, 14(S2), 95-112.
- Levitas, E. F., McFadyen, M., & Loree, D. (2006). Survival and the introduction of new technology: a patent analysis in the integrated circuit industry. *Journal of Engineering and Technology Management*, 23(3), 182-201.
- Li, G.-C., Lai, L., D'Amour, A., Doolin, D., Sun, Y., Torvik, V., Yu, A., Fleming L. (2014). Disambiguation and co-authorship networks of the U.S. patent inventor database (1975–2010). *Research Policy*, 43, 941-955.
- Linden, G., & Somaya, D. (2003). System-on-a-chip integration in the semiconductor industry: industry structure and firm strategies. *Industrial and Corporate Change*, 12(3), 545-576.
- Lu, C. (2007). Moving up or moving out? A unified theory of R&D, FDI, and trade. *Journal of International Economics*, 71(2), 324–343.
- Luo, J., Olechowski, A. L., & Magee, C. L. (2012). Technology-based design and sustainable economic growth. *Technovation*.
- Malerba, F., Nelson, R., Orsenigo, L., & Winter, S. (1999). 'History-friendly' models of industry evolution: the computer industry. *Industrial and Corporate Change*, 8(1), 3-40.
- Malerba, F., & Orsenigo, L. (1993). Technological regimes and firm behaviour. *Industrial and corporate change*, 2(1), 45-71.
- Malerba, F., & Orsenigo, L. (1995). Schumpeterian patterns of innovation. *Cambridge Journal of Economics*, 19(1), 47-65.
- Malerba, F., & Orsenigo, L. (1997). Technological regimes and sectoral patterns of innovative activities. *Industrial and corporate change*, 6(1), 83-118.
- Malerba, F., Orsenigo, L., & Peretto, P. (1997). Persistence of innovative activities, sectoral patterns of innovation and international technological specialization. *International Journal of Industrial Organization*, 15(6), 801-826.
- Magee, C. L. (2010). The role of materials innovation in overall technological development. *Journal of Management*, 62(3), 20-24.
- March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization science*, 2(1), 71-87.
- Martinelli A. (2008), Technological Trajectories and Industry Evolution: The case of Telecom Switching Industry, Paper presented at the 25th DRUID Summer Conference, Copenhagen.
- Martinelli A. (2009), Market Dynamics and Technological Competences in Ologopolistic Sectors: The Case of Telecom Switches, Paper presented at the Druid Summer Conference 2009.
- Martinelli A., Nomaler Ö. (2014), Measuring knowledge persistence: A genetic approach to patent citation networks. *Journal of Evolutionary Economics*, 24(3), 632-52.
- Mathews, J. A., & Cho, D. S. (1999). Combinative capabilities and organizational learning in latecomer firms: The case of the Korean semiconductor industry. *Journal of World Business*, 34(2), 139-156.

- Metcalf, S. (2014). Capitalism and evolution. *Journal of Evolutionary Economics*, 24(1), 11-34.
- Moore G.E. (1965), "Cramming more components onto integrated circuits", Electronics Magazine, p.4.
- Murmann, J. P., & Frenken, K. (2006). Toward a systematic framework for research on dominant designs, technological innovations, and industrial change. *Research Policy*, 35(7), 925-952.
- Neffke, F., & Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal*, 34(3), 297-316.
- Newman, M. E. (2004). Fast algorithm for detecting community structure in networks. *Physical review E*, 69(6), 066133.
- Newman, M. E. (2005). A measure of betweenness centrality based on random walks. *Social networks*, 27(1), 39-54.
- O'Reilly III, C. A., & Tushman, M. L. (2008). Ambidexterity as a dynamic capability: Resolving the innovator's dilemma. *Research in organizational behavior*, 28, 185-206.
- Parpala M. (2014). The US Semiconductor industry: a key contributor to U.S. economic growth. SIA (Semiconductor Industry Association) August 2014 white paper.
- Perez, C. (1988). New technologies and development. In: Freeman, C., Lundvall, B. (Eds.), *Small Countries Facing the Technological Revolution*. Pinter Publishers, London.
- Perez, C., & Soete, L. (1988). Catching up in technology: entry barriers and windows of opportunity. In: Dosi, G. et al. (eds) *Technical change and economic theory*. Pinter, London, pp 458-479.
- Peters, B. (2009). Persistence of innovation: stylised facts and panel data evidence. *The Journal of Technology Transfer*, 34(2), 226-243.
- Raymond, W., Mohnen, P., Palm, F., & van der Loeff, S. S. (2010). Persistence of innovation in Dutch manufacturing: Is it spurious?. *The Review of Economics and Statistics*, 92(3), 495-504.
- Raisch, S., Birkinshaw, J., Probst, G., & Tushman, M. L. (2009). Organizational ambidexterity: Balancing exploitation and exploration for sustained performance. *Organization Science*, 20(4), 685-695.
- Reitzig, M. (2003). What determines patent value?: Insights from the semiconductor industry. *Research Policy*, 32(1), 13-26.
- Rodriguez, G., & Elo, I. (2003). Intra-class correlation in random-effects models for binary data. *The Stata Journal*, 3(1), 32-46.
- Roper, S., & Hewitt-Dundas, N. (2008). Innovation persistence: Survey and case-study evidence. *Research Policy*, 37(1), 149-162.
- Rosenbloom, R. S., & Christensen, C. M. (1994). Technological discontinuities, organizational capabilities, and strategic commitments. *Industrial and corporate change*, 3(3), 655-685.
- Salzano, M., & Colander, D. (2007). *Complexity hints for economic policy*. Springer.
- Sapienza, H. J., Autio, E., George, G., & Zahra, S. A. (2006). A capabilities perspective on the effects of early internationalization on firm survival and growth. *Academy of management review*, 31(4), 914-933.
- Saviotti, P. P., & Pyka, A. (2004). Economic development by the creation of new sectors. *Journal of evolutionary economics*, 14(1), 1-35.
- Schmookler, J. (1962). Changes in industry and in the state of knowledge as determinants of industrial invention. In Groves H.M. (1962). *The rate and direction of inventive activity: Economic and social factors* (pp. 195-232). Princeton University Press.
- Shibata N., Kajikawa Y., Takeda Y., Matsushima K. (2008). Detecting emerging research fronts based on topological measures in citation networks of scientific publications. *Technovation* 28, pp. 758-775.
- Silverberg, G., Dosi, G., & Orsenigo, L. (1988). Innovation, diversity and diffusion: a self-organisation model. *The Economic Journal*, 1032-1054.
- Silverberg, G., & Verspagen, B. (2005). A percolation model of innovation in complex technology spaces. *Journal of Economic Dynamics and Control*, 29(1), 225-244.

- Silverberg, G., & Verspagen, B. (2007). The size distribution of innovations revisited: an application of extreme value statistics to citation and value measures of patent significance. *Journal of Econometrics*, 139(2), 318-339.
- Soete, L. (2000). Technology, Globalisation and Employment: Analytical and Policy Challenges. In *Education and training in a knowledge-based economy*. Ed. by Hans Heijke and Joan Muyksen, Chapter 11, 279-299.
- Soete L. (1987), The impact of technological innovation on international trade patterns: the evidence reconsidered, *Research Policy*, Vol.16(2), pp. 101-130.
- Suarez, F. F., & Utterback, J. M. (1995). Dominant designs and the survival of firms. *Strategic management journal*, 16(6), 415-430.
- Teece, D., & Pisano, G. (1994). The dynamic capabilities of firms: an introduction. *Industrial and corporate change*, 3(3), 537-556.
- Teece, D. J., Rumelt, R., Dosi, G., & Winter, S. (1994). Understanding corporate coherence: Theory and evidence. *Journal of Economic Behavior & Organization*, 23(1), 1-30.
- Thurner, S., Klimek, P., & Hanel, R. (2010). Schumpeterian economic dynamics as a quantifiable model of evolution. *New Journal of Physics*, 12(7), 075029.
- Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *The Rand Journal of Economics*, 172-187.
- USPTO (2014), Manual of Patent Examining Procedure, March 2014 Edition, retrieved online at <http://www.uspto.gov/web/offices/pac/mpep/>
- Utterback, J. M. (1996). *Mastering the dynamics of innovation*. Harvard Business Press.
- Utterback, J. M., & Abernathy, W. J. (1975). A dynamic model of process and product innovation. *Omega*, 3(6), 639-656.
- Uzzi B., Mukherjee S., Stringer M., Jones B. (2013). Atypical Combinations and Scientific Impact. *Science* 342(6157), 468-472
- Vernon R. (1966), "International Investments and International Trade in the Product Cycle", *The Quarterly Journal of Economics*, Vol. 80, No. 2, pp. 190-207.
- Verspagen B. (1991), "A New Empirical Approach to Catching Up and Falling Behind", *Structural Change and Economic Dynamics*, Vol. 2, pp. 359-380.
- Verspagen B. (2007), "Mapping technological trajectories as patent citation networks: A study on the history of fuel cell research", *Advances in Complex Systems*, Vol. 10 (1), pp. 93-115.
- Vértesy, D. (2011). *Interrupted Innovation Emerging Economies in the Structure of the Global Aerospace Industry*. Datawyse/Universitaire Pers Maastricht.
- Williams, R. (2012). Using the margins command to estimate and interpret adjusted predictions and marginal effects. *Stata Journal*, 12(2), 308.
- Winter, S. G. (2003). Understanding dynamic capabilities. *Strategic management journal*, 24(10), 991-995.
- Xu, S. (2014). Balancing the Two Knowledge Dimensions in Innovation Efforts: An Empirical Examination among Pharmaceutical Firms. *Journal of Product Innovation Management*.
- Yinug (2015). *Made in America: the facts about Semiconductor manufacturing*. SIA (Semiconductor Industry Association) August 2015 white paper.





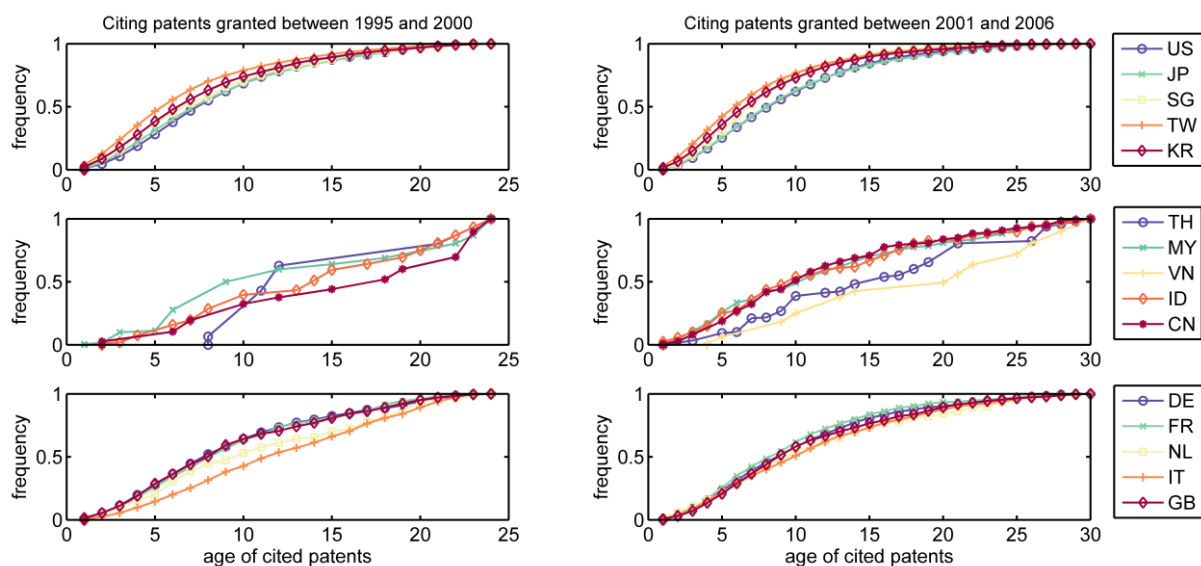
# 8 APPENDICES

## APPENDIX TO CHAPTER 2

### **A.2.1 Normalization of the number of between-firms, within-country citations by randomizing citation networks**

To assess the extent to which patenting organizations are building upon foreign or locally invented technologies we need to have an expectation of the number of citations that falls between countries and one of the number of within-country citations. One way of building an expectation for these values is to shuffle citations randomly and then compute the expected value and standard deviation of the number of inter- and intra-country citations. When doing so it is crucial to create random realizations of the system that are plausible given what we observe in the real world. There are a number of phenomena that must be preserved in the randomized networks. First, the distribution of citations made and received by patents are strongly skewed (most of their cumulative distribution function follows a power-law). Second, the probability of receiving a citation is a function of the age of the patent. Inventions need time to be recognized and old inventions are usually considered obsolete. Therefore, the probability of receiving a citation initially increase with age, peak at around 4 years and then rapidly decrease, as shown by Jaffe and Trajtenberg (2002) and Hall et al. (2001). Finally, patenting organizations have a marked tendency to cite their previous patents. This is because learning is path-dependent and many technical improvements are incremental in nature. Consequently, firms tend to improve upon technical solutions that they have developed in the past. Therefore, to generate plausible random realization of the real world, the randomization process must reproduce the in- and out-degree distributions, the typical aging function of citations and the share of self-citations. The easier way of doing so is to swap inter-firm citations for each citing-cited year pair. For each citing-cited year pair our algorithm swap all inter-firm citations. This preserves both the in- and out-degree of each patent (and therefore their distribution) and the lag between the citing and the cited grant year of each citation. The number of inter-firm between-country citations and its counterpart for within-country citations then becomes random variables. If we repeat the process 1000 times we obtain a mean and a standard deviation for these statistics. The mean approximate their true expected value. We can then compute the z-scores for the number of inter-firm between- and within-country citations reported in Chapter 2. Note that, under certain conditions the expected value and the standard deviation can be predicted by the sum of hypergeometric random variables. This is discussed in details in another paper of the same author of this thesis (Alstott et al., 2015).

## A.2.2 Empirical cumulative distribution functions of the probability of citing a patent as a function of its age by citing country



**Figure 38: Empirical cumulative distribution functions of the probability of citing a patent as a function of its age by citing country.**

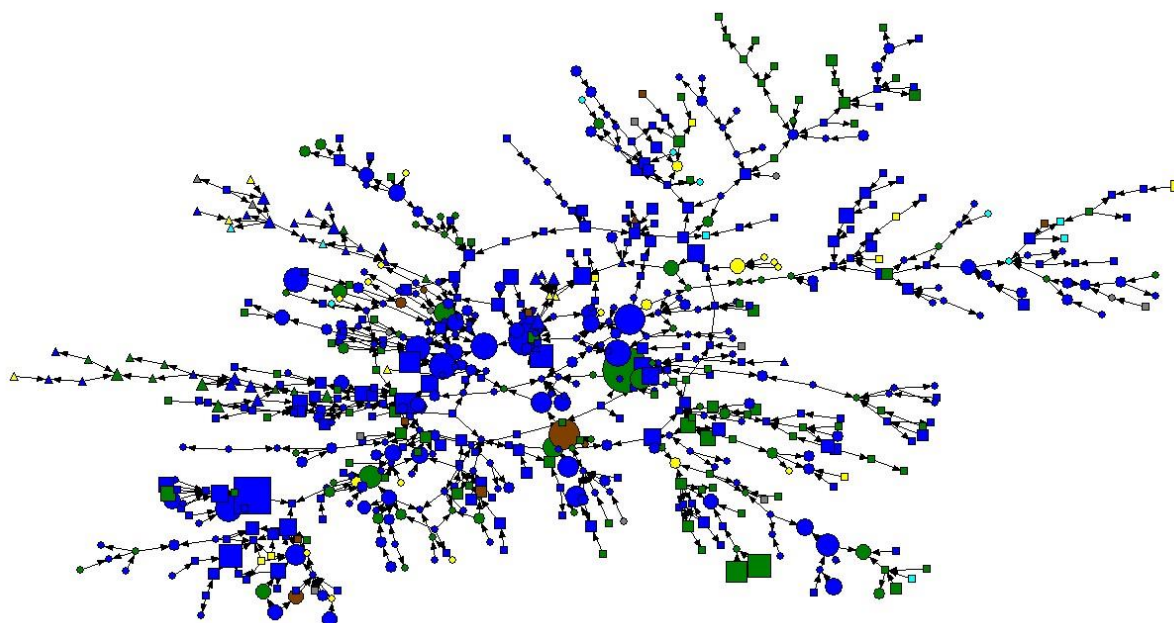
## APPENDIX TO CHAPTER 3

### A.3.1 Visual analysis of the main component of the Network of Main Paths

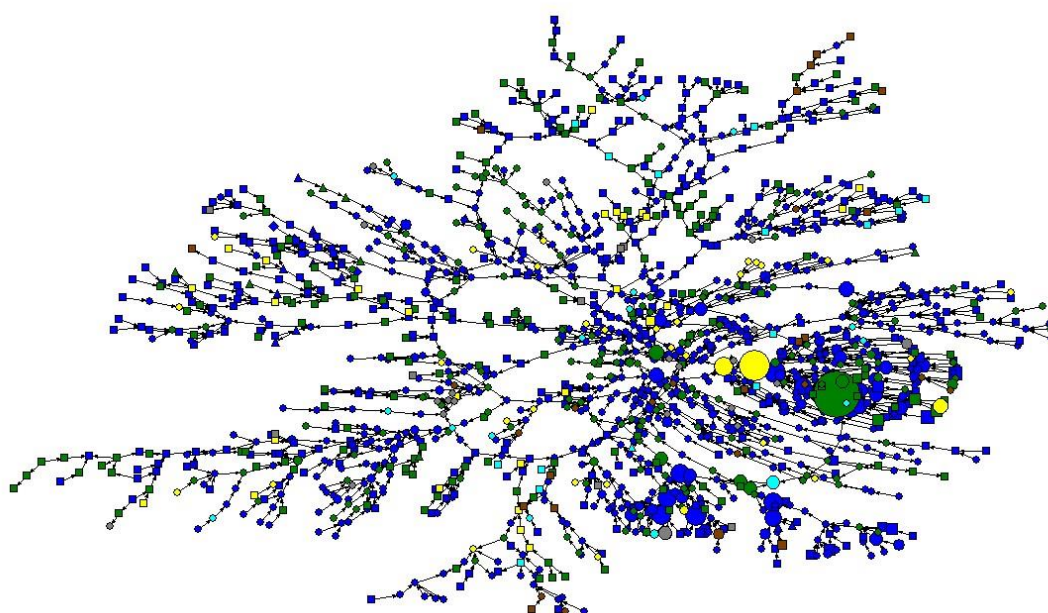
We inspect the composition of the main component of the NMPs, by plotting its visual appearance for each of the six periods considered. These graphs are reported in figures from Figure 30 to Figure 44. Table 29 reports the legend of nodes' colour, shape and size, to help interpreting the graphs. We can group the insights provided by the visual analysis of the NMPs along two themes: technological classes and countries. If we look at shapes of the nodes we can notice that patents belonging to class 438 (process innovation) have always been the majority, especially among the highly connected nodes (those bigger in size). However, from 1976-1995 they virtually overthrown other classes, except for a few largely connected US and Japanese patents assigned to class 257 (product innovation), which show up in the main component of the NMPs in periods 1976-2000 and 1976-2006. This reveals an increasing innovative effort devoted to engineering problems related to process innovation. Indeed, in the first three periods, it was still possible to find well connected patents related to product innovation (class 257). Furthermore, in periods 1976-1985 and 1976-1990, there were also areas of the network of main paths almost entirely populated by patents belonging to class 326. These areas almost disappear from the main component of the NMPs in the last three periods. The second important insight comes from comparing location of assignees from different countries. The NMPs in the first three periods is dominated by US and Japanese patents, with the presence of a few patents granted to German assignees. In the period 1976-1995 the reader can notice the appearance of many red nodes (standing for Taiwanese assignees), and a few lilac ones (South Korean assignees) located in crucial parts of the network (on the backbone and in the key junctions). This becomes much more visible in the period 1976-2000, when the size of Taiwanese and Korean patents increases, revealing an increased centrality of their technological solutions. Interestingly, in the final period 1976-2006, US patents seem to regain the lead in terms of centrality.

**Table 29: Legend of nodes' colour, shape and size**

| Colour | Assignee's country | Shape         | U.S. main technology class         |
|--------|--------------------|---------------|------------------------------------|
| Blue   | U.S.A.             | Circle        | 438 – Process                      |
| Green  | Japan              | Square        | 257 – Product                      |
| Yellow | Germany            | Up-triangle   | 326 - Programmability              |
| Cyan   | Italy and France   | Diamond       | 716 - Design                       |
| Purple | S.Korea            | Down-triangle | 505 – Material                     |
| Red    | Taiwan             | <b>Size</b>   | <b>Proportional to PathC Index</b> |
| Black  | Singapore          |               |                                    |
| Orange | Netherlands        |               |                                    |
| Grey   | No assignee        |               |                                    |
| Brown  | Others             |               |                                    |



**Figure 39: Main component of the NMPs 1976-1980**



**Figure 40: Main component of the NMPs 1976-1985**

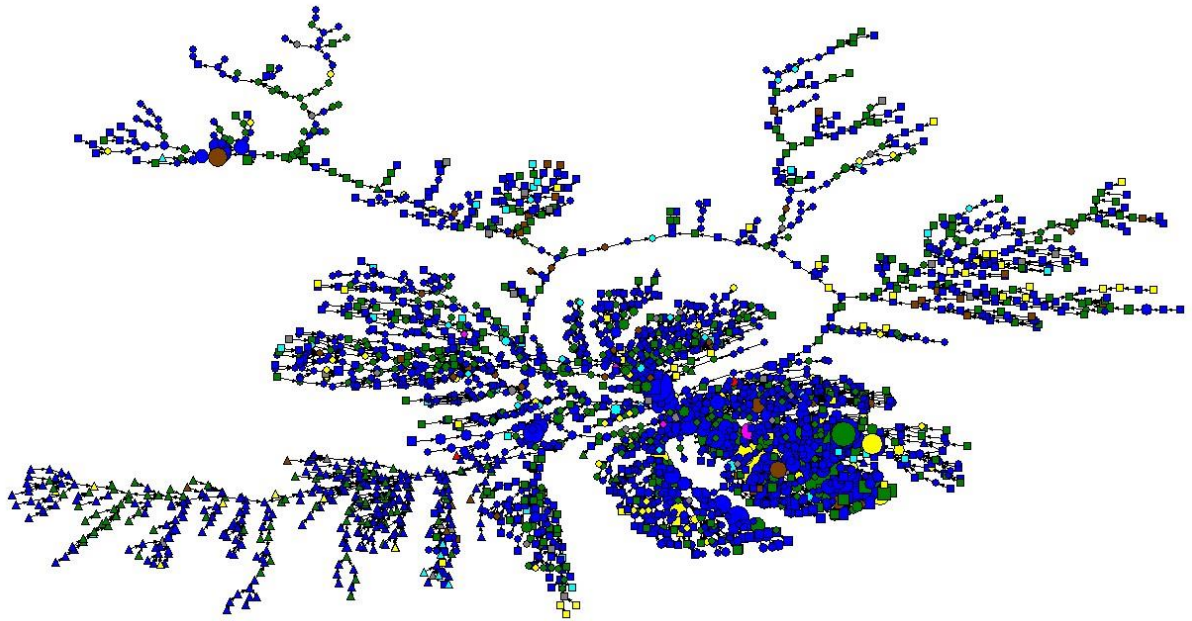


Figure 41: Main component of the NMPs 1976-1990

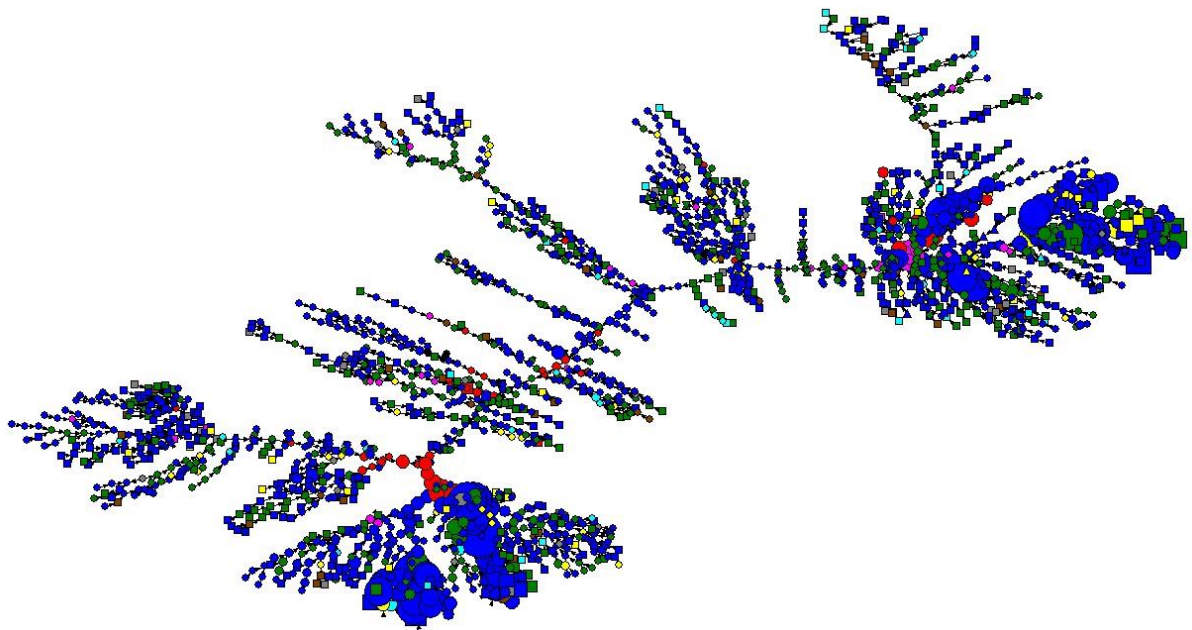


Figure 42: Main component of the NMPs 1976-1995



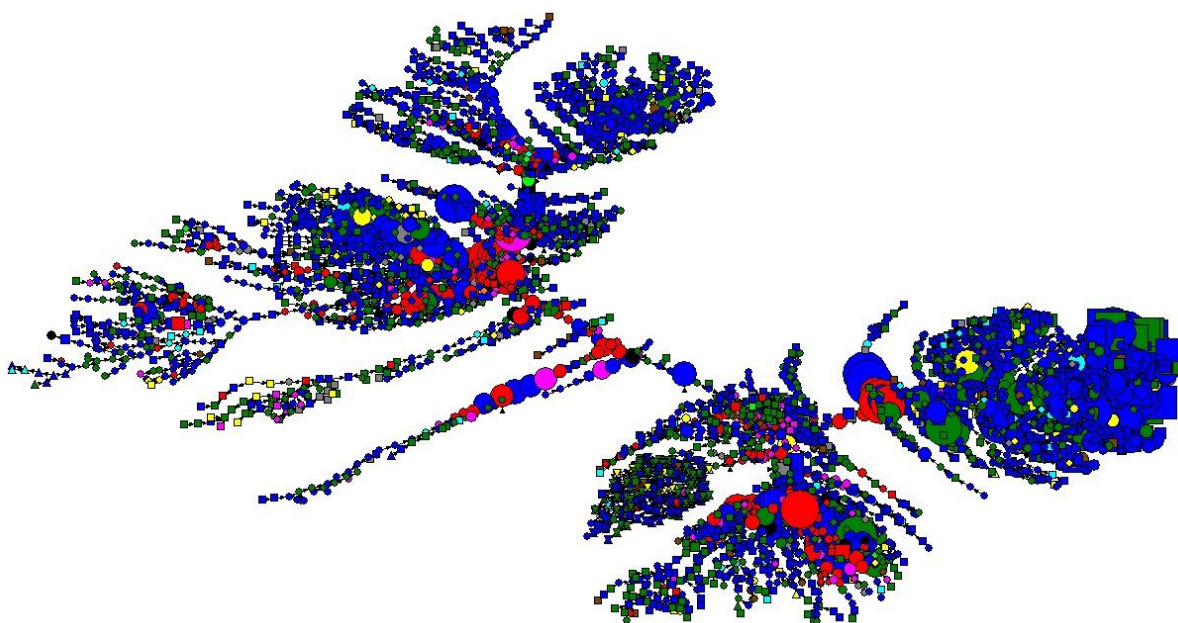


Figure 43: Main component of the NMPs 1976-2000

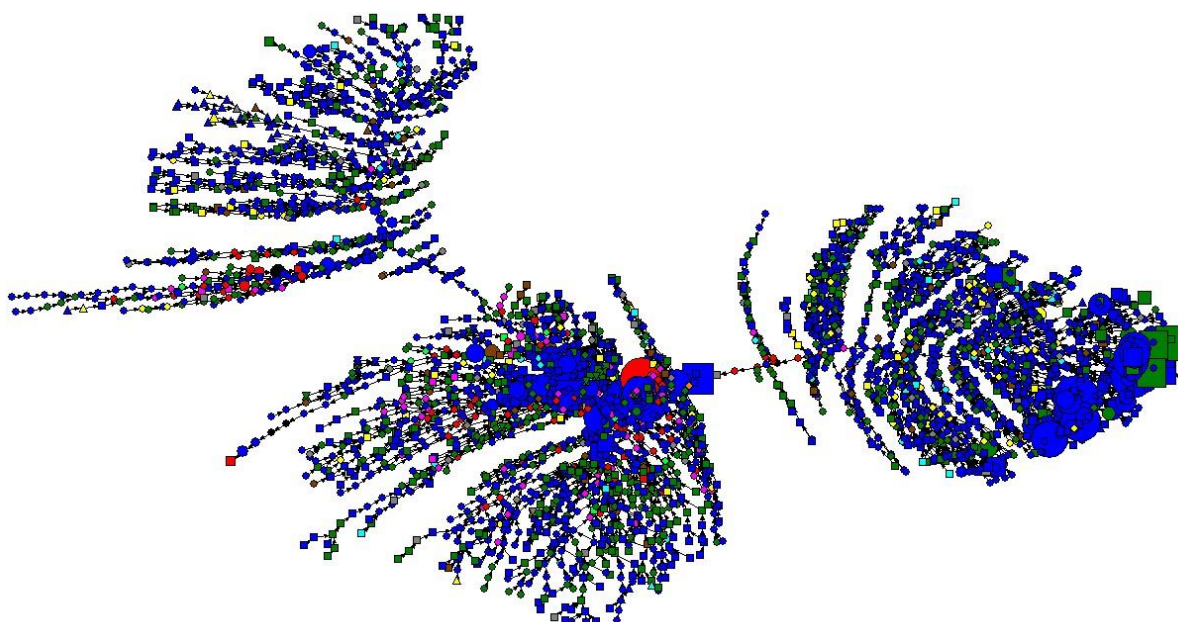
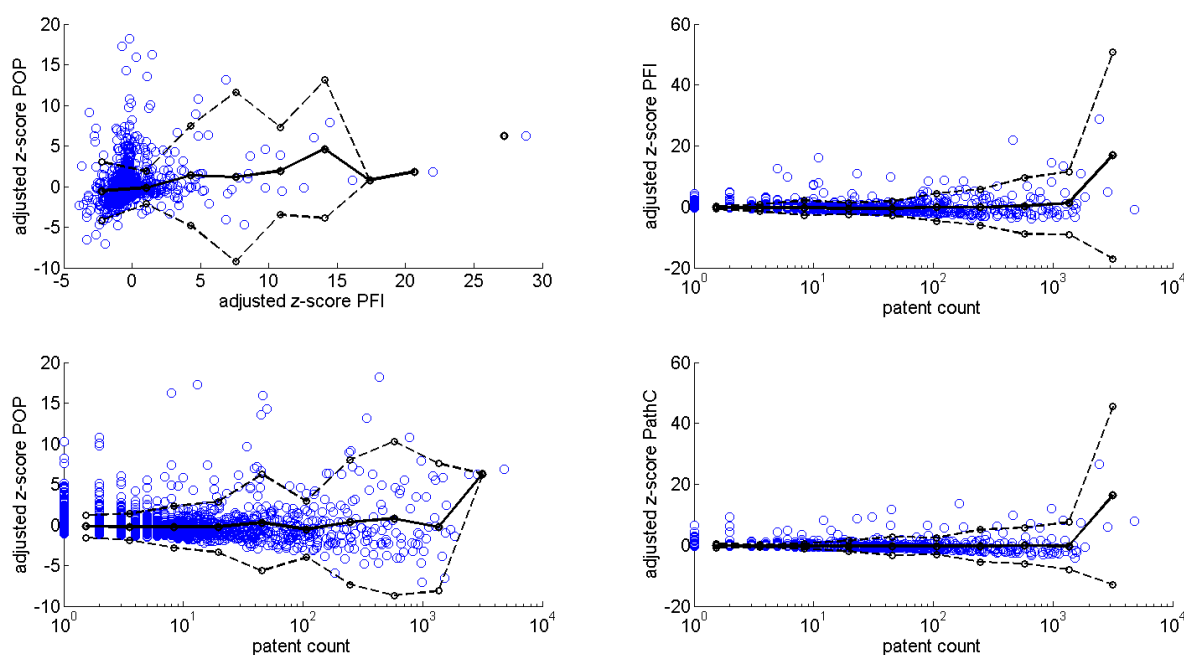


Figure 44: Main component of the NMPs 1976-2006

### A.3.2 Statistical analysis of the relationship between PFI, POP, PathC and patent count

We report here the statistical analysis of the relationship between the path-following (PFI) and the power-of-pull (POP) indices. We also show how the bootstrapping technique used to clean the POP, PFI and Path-centrality (PathC) indices from firm size effects, effectively reveals the lack of a true relationship between firms size and innovation strategies.

No curve could be meaningfully fitted on the data (the  $R^2$  for the best fit was on average 0.1). The solid line indicates the mean of the y variable for each bin of the x variable. Dashed lines are drawn above and below two standard deviations from the mean. When the variable on the horizontal axis was patent count, logarithmic binning was used to reduce noise due to scarcity of observations in the right tail of the distribution.

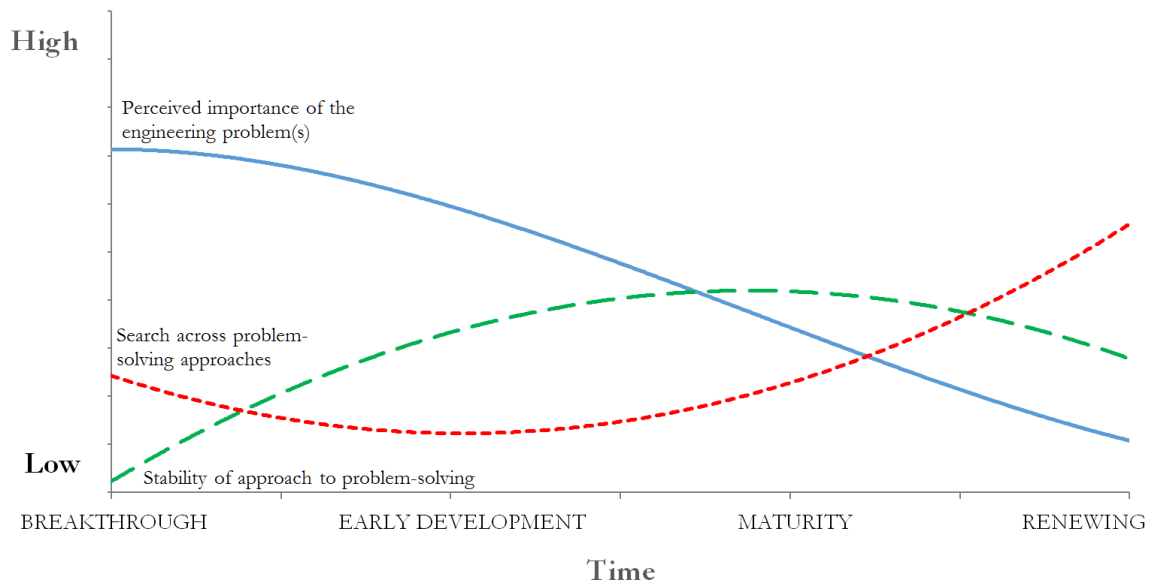


**Figure 45: Statistical relationship between PFI, POP, PathC indices and firms' patent count**

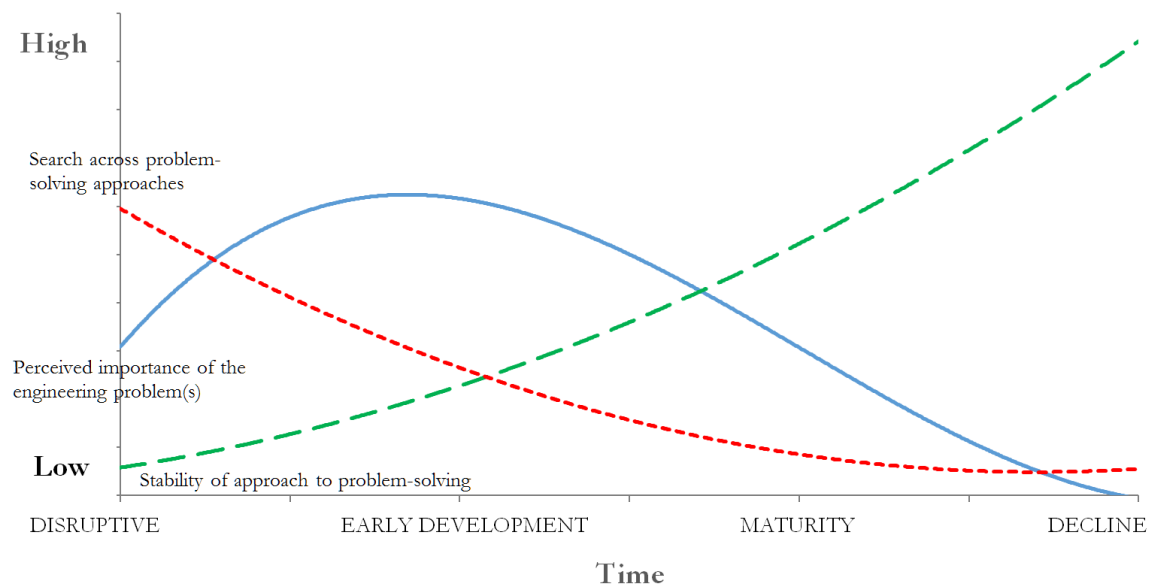


## APPENDIX TO CHAPTER 4

### A.4.1 Alternative beginning and end of the archetypal life-cycle of technology domains



**Figure 46: Archetypal life-cycle of a given technology domain with resistance to decline**



**Figure 47: Archetypal life-cycle of a given technology domain starting with a disruption**

#### A.4.2 Plots of the main component of the network of main paths (NMPs)

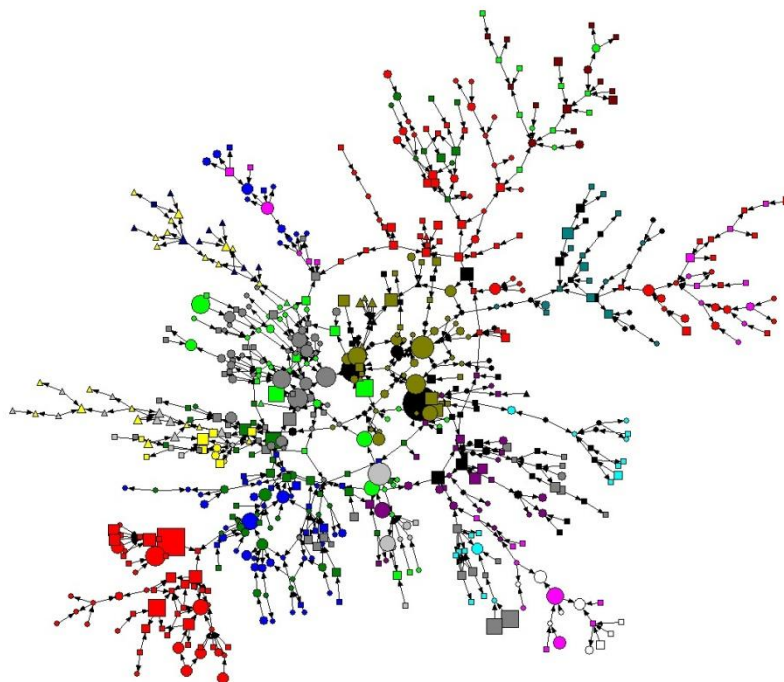


Figure 48: The space of technology domains between 1976 and 1980

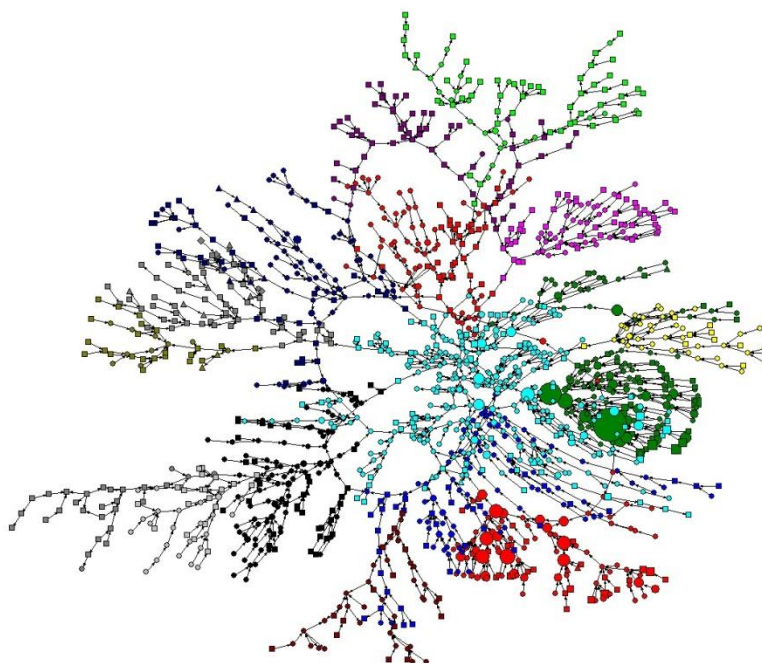
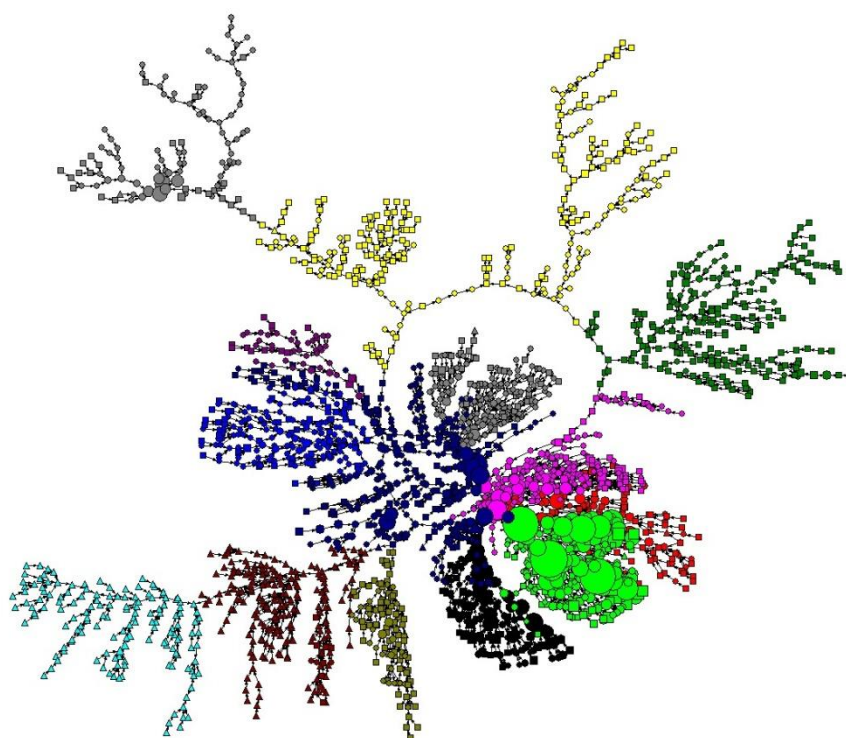
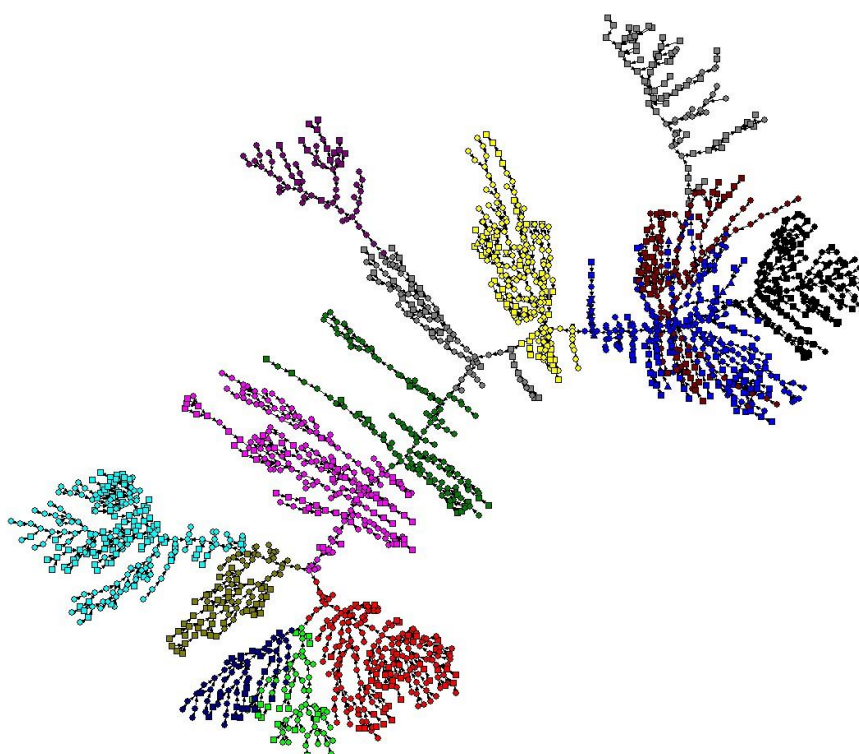


Figure 49: The space of technology domains between 1976 and 1985



**Figure 50: The space of technology domains between 1976 and 1990**



**Figure 51: The space of technology domains between 1976 and 1995**

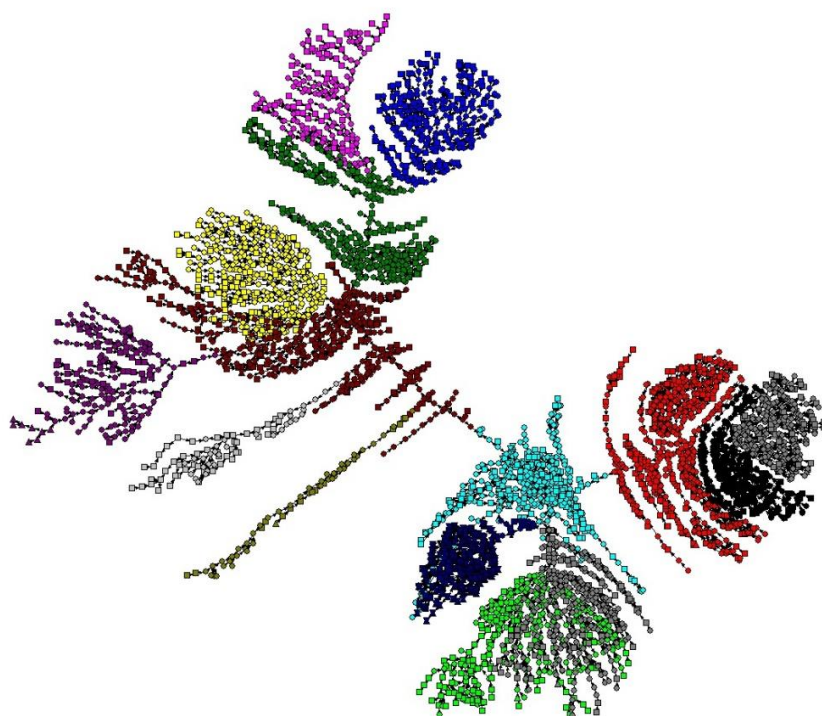


Figure 52: The space of technology domains between 1976 and 2000

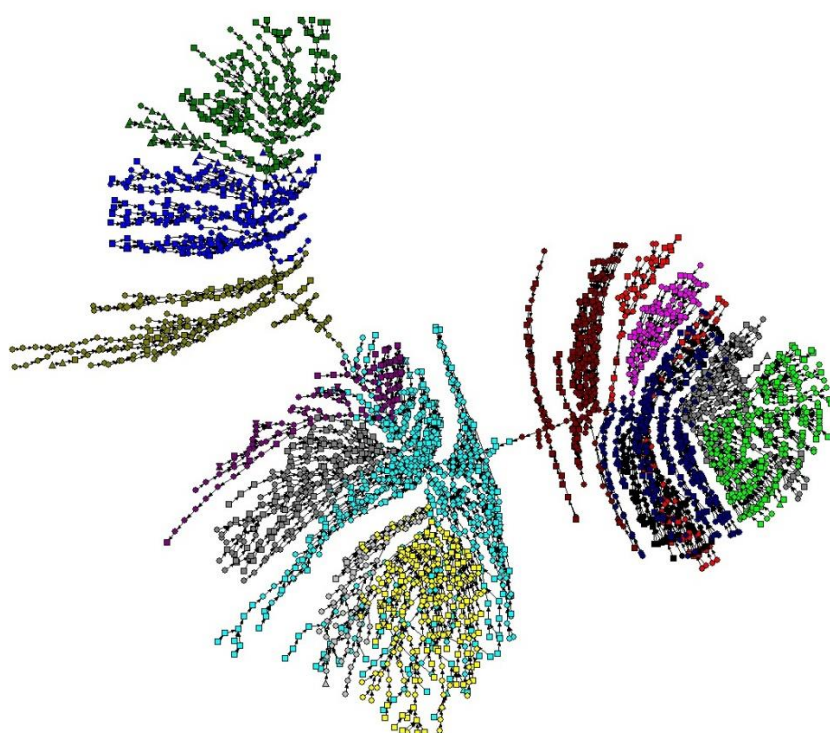


Figure 53: The space of technology domains between 1976 and 2006

### A.4.3 Newman's community detection algorithm

To identify technology domains we used Newman's algorithm (Newman, 2004). The algorithm maximizes modularity  $Q$ , which is defined as follows:

$$Q = \sum_i (e_{ii} - a_i^2) \quad (\text{A.4.1})$$

Where  $e_{ii}$  is the fraction of edges falling within community  $i$  and  $a_i^2$  is equal to the squared sum of edges falling between communities, as  $a_i = \sum_j e_{ij}$ . Newman (2004) explains that modularity  $Q$  can be also calculated as the fraction of edges that fall within communities, minus the expected value of the same quantity if edges fall at random without regard for the community structure. The author highlights that if a particular division gives no more within-community edges than would be expected by random chance modularity  $Q$  would be equal to zero. This approach allows optimizing modularity  $Q$  without the need to try all possible partition combinations (which would take an amount of time exponential to the number of nodes in the network). The optimization approach starts from the worse possible combination. It then begins an iterative aggregation process that stops when the increase of modularity becomes negative. Obviously, as explained by Newman (2004), since the joining of a pair of communities between which there are no edges at all can never result in an increase in  $Q$ , one needs only consider those pairs between which there are edges. Then the change in  $Q$  upon joining two communities is given by:

$$\Delta Q = e_{ij} + e_{ji} - 2a_i a_j = 2(e_{ij} - a_i a_j) \quad (\text{A.4.2})$$

One possible drawback of Newman's algorithm is that it is not specifically thought for citations network, which have the peculiarity to be acyclical directed graphs. Yet, symmetrizing the adjacency matrix makes citations a univocal measure of relatedness from patent to patent. This allows using the algorithm. The second possible limitation consists in the fact that a real-world citation networks are sparser than the random counterparts that are used as benchmark to maximize modularity. This is due to the well-known shape of the distribution of citation-lags for patent networks. Jaffe and Trajtenberg (2002) showed that citations received by the average patent peaks after 3-4 years and then sharply decline. This is because constant streams of technical improvements make older patents irrelevant for the legal definition of the prior-art. Potentially this bias can identify communities on the network purely based on their age structure of patent citations, without considering the true relationship of similarity that might exist with older patents. To assess the strength of this bias we analysed the age structure of the communities (i.e. technology domains) identified by the algorithm. Results are shown in Figure 54. The domains' density of patents for each time cohort is shown by mean of a density plot where darker colours represent higher density. We can clearly see that a few domains that are time dependent are visible only in the last period. Since there are few examples we cannot discard the possibility that these domains are indeed declining, i.e. their underlying engineering problems failed to attract further attention. The fact that age dependent communities are very rare proves that the potential bias in the algorithm does not affect the quality and validity of our results.



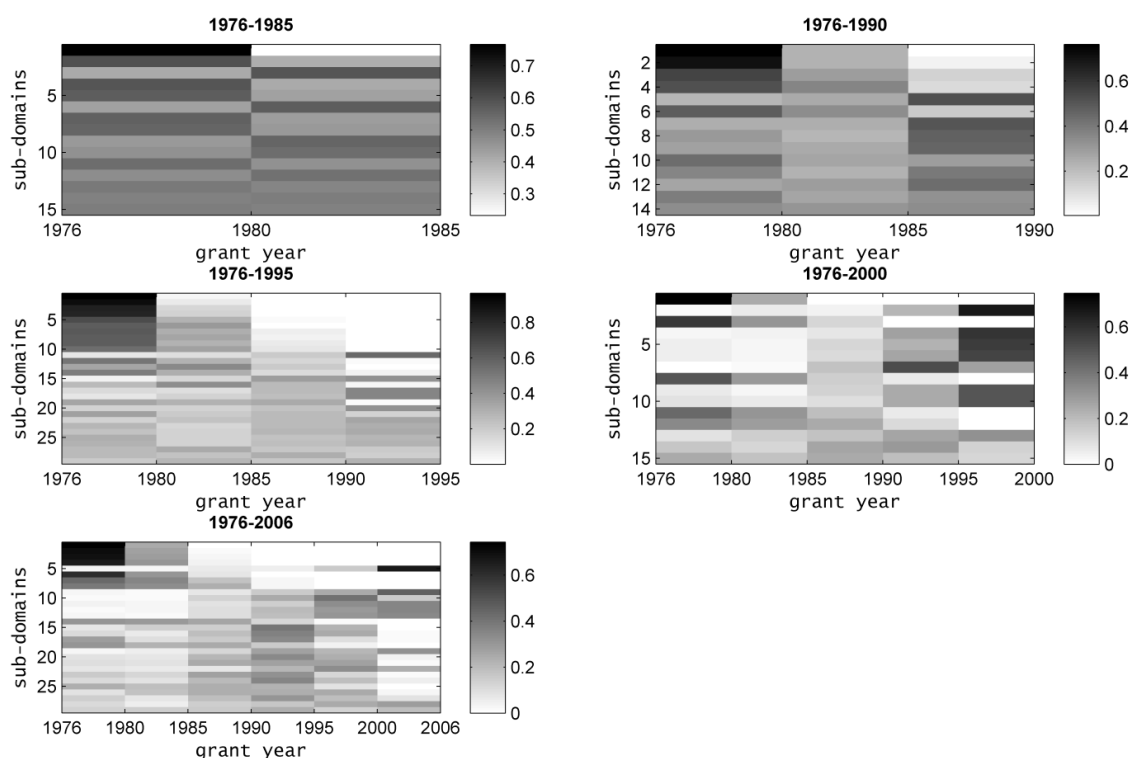


Figure 54: Age structure of technology domains

#### A.4.4 Anderson-Darling test results

In this section, we report the results of the non-parametric two-sample Anderson-Darling test for statistical difference.

| Andersen-Darling test | Ninc | Nnew | Test result                     | p-value |
|-----------------------|------|------|---------------------------------|---------|
| Breakthrough          | 201  | 213  | H1:Fnew(SRTA) $\neq$ Finc(SRTA) | 0.0000  |
| Disruptive            | 381  | 368  | H1:Fnew(SRTA) $\neq$ Finc(SRTA) | 0.0039  |
| Early growth          | 266  | 287  | H1:Fnew(SRTA) $\neq$ Finc(SRTA) | 0.0000  |
| Mature                | 381  | 368  | H1:Fnew(SRTA) $\neq$ Finc(SRTA) | 0.0000  |
| Renewing              | 336  | 314  | H1:Fnew(SRTA) $\neq$ Finc(SRTA) | 0.0000  |
| Exhausting            | 336  | 314  | H0: Fnew(SRTA) = Finc(SRTA)     | 0.8460  |

| BREAKTHROUGH     |            |            |                  |                     |
|------------------|------------|------------|------------------|---------------------|
|                  | US new inn | JP new inn | KR/TW/SG new inn | All incumbent innov |
|                  | p-value    | p-value    | p-value          | p-value             |
| US new inn       | --         | 0.8380     | 0.1495           | 0.0000              |
| JP new inn       | 0.8380     | --         | 0.3939           | 0.0009              |
| KR/TW/SG new inn | 0.1495     | 0.3939     | --               | 0.1050              |
| EARLY GROWTH     |            |            |                  |                     |
|                  | p-value    | p-value    | p-value          | p-value             |
| US new inn       | --         | 0.0001     | 0.3999           | 0.0027              |
| JP new inn       | 0.0001     | --         | 0.0111           | 0.0000              |
| KR/TW/SG new inn | 0.3999     | 0.0111     | --               | 0.0990              |
| RENEWING         |            |            |                  |                     |
|                  | p-value    | p-value    | p-value          | p-value             |
| US new inn       | --         | 0.7154     | 0.0016           | 0.0000              |
| JP new inn       | 0.7154     | --         | 0.0009           | 0.0000              |
| KR/TW/SG new inn | 0.0016     | 0.0009     | --               | 0.7093              |

| DISRUPTIVE       |            |            |                  |                     |
|------------------|------------|------------|------------------|---------------------|
|                  | US new inn | JP new inn | KR/TW/SG new inn | All incumbent innov |
|                  | p-value    | p-value    | p-value          | p-value             |
| US new inn       | --         | 0.0000     | 0.1301           | 0.1856              |
| JP new inn       | 0.0000     | --         | 0.0000           | 0.0000              |
| KR/TW/SG new inn | 0.1301     | 0.0000     | --               | 0.0545              |
| MATURE           |            |            |                  |                     |
|                  | p-value    | p-value    | p-value          | p-value             |
| US new inn       | --         | 0.8759     | 0.8773           | 0.0001              |
| JP new inn       | 0.8759     | --         | 0.8817           | 0.0037              |
| KR/TW/SG new inn | 0.8773     | 0.8817     | --               | 0.0325              |
| EXHAUSTING       |            |            |                  |                     |
|                  | p-value    | p-value    | p-value          | p-value             |
| US new inn       | --         | 0.5527     | 0.6107           | 0.8625              |
| JP new inn       | 0.5527     | --         | 0.1360           | 0.5487              |
| KR/TW/SG new inn | 0.6107     | 0.1360     | --               | 0.4532              |

#### **A.4.5 Topic analysis of the main technology domains of the Semiconductor Industry between 2001 and 2006**

In this section, we report the title of the most central patents within each technology domain identified by the Newman's modularity maximization algorithm.

**Table 30: Topic analysis of the main technology domains of the Semiconductor industry between 2001 and 2006**

| Patent  | grant year | Cluster | LC stage   | NMP comp | PathC | Title   | Assignee          |
|---------|------------|---------|------------|----------|-------|---|-------------------|
| 6451641 | 2002       | 3       | Exhausting | 1        | 13.40 | Non-reducing process for deposition of polysilicon gate electrode over high-K gate dielectric material                                      | AMD               |
| 6297539 | 2001       | 3       | Exhausting | 1        | 2.49  | Zirconium or hafnium oxide doped with calcium, strontium, aluminum, lanthanum, yttrium, or scandium   | SHARP             |
| 6407435 | 2002       | 3       | Exhausting | 1        | 2.16  | Because the layers reduce the effects of crystalline structures within individual layers, the overall tunneling current is reduced.         | SHARP             |
| 6207589 | 2001       | 3       | Exhausting | 1        | 0.20  | Method of forming a doped metal oxide dielectric film   | SHARP             |
| 6297107 | 2001       | 6       | Exhausting | 1        | 12.96 | High dielectric constant materials as gate dielectrics  | AMD               |
| 6200865 | 2001       | 6       | Exhausting | 1        | 7.57  | Use of sacrificial dielectric structure to form semiconductor device with a self-aligned threshold adjust and overlying low-resistance gate | AMD               |
| 6391785 | 2002       | 7       | Mature     | 1        | 1.85  | Method for bottomless deposition of barrier layers in integrated circuit metallization schemes  | ASM/IMEC          |
| 6184128 | 2001       | 7       | Mature     | 1        | 2.61  | Method using a thin resist mask for dual damascene stop layer etch  | AMD               |
| 6468924 | 2002       | 7       | Mature     | 1        | 0.73  | Methods of forming thin films by atomic layer deposition  | SAMSUNG           |
| 6750066 | 2004       | 7       | Mature     | 1        | 2.32  | Precision high-K intergate dielectric layer   | AMD               |
| 6534395 | 2003       | 7       | Mature     | 1        | 0.83  | Method of forming graded thin films using alternating pulses of vapor phase reactants   | ASM               |
| 6424001 | 2002       | 9       | Renewing   | 1        | 1.34  | Flash memory with ultra thin vertical body transistors  | MICRON            |
| 6639268 | 2003       | 9       | Renewing   | 1        | 0.61  | Flash memory with ultra thin vertical body transistors  | MICRON            |
| 6680508 | 2004       | 9       | Renewing   | 1        | 1.36  | Vertical floating gate transistor   | MICRON            |
| 6903367 | 2005       | 9       | Renewing   | 1        | 0.32  | Programmable memory address and decode circuits with vertical body transistors  | MICRON            |
| 6979857 | 2005       | 9       | Renewing   | 1        | 0.32  | Apparatus and method for split gate NROM memory   | MICRON            |
| 6303523 | 2001       | 11      | Exhausting | 1        | 1.02  | Plasma processes for depositing low dielectric constant films   | APPLIED MATERIALS |
| 6410462 | 2002       | 11      | Exhausting | 1        | 0.57  | Method of making low-K carbon doped silicon oxide   | SHARP             |
| 6287990 | 2001       | 11      | Exhausting | 1        | 0.90  | CVD plasma assisted low dielectric constant films   | APPLIED MATERIALS |
| 6534397 | 2003       | 12      | Disruptive | 1        | 1.65  | Pre-treatment of low-k dielectric for prevention of photoresist poisoning   | AMD               |
| 6656837 | 2003       | 12      | Disruptive | 1        | 1.78  | Method of eliminating photoresist poisoning in damascene applications   | APPLIED MATERIALS |
| 6406994 | 2002       | 12      | Disruptive | 1        | 1.79  | Triple-layered low dielectric constant dielectric dual damascene approach   | CHARTERED         |
| 6593247 | 2003       | 12      | Disruptive | 1        | 0.87  | Method of depositing low k films using an oxidizing plasma  | APPLIED MATERIALS |
| 6784119 | 2004       | 12      | Disruptive | 1        | 0.37  | Method of decreasing the K value in SIOC layer deposited by chemical vapor deposition   | APPLIED MATERIALS |



| Patent  | grant<br>year | Cluster | LC stage     | NMP<br>comp | PathC | Title  | Assignee           |
|---------|---------------|---------|--------------|-------------|-------|--|--------------------|
| 6979855 | 2005          | 15      | Renewing     | 1           | 8.45  | High-quality praseodymium gate dielectrics   | MICRON             |
| 7045430 | 2006          | 15      | Renewing     | 1           | 3.68  | Atomic layer-deposited LaAlO3 films for gate dielectrics   | MICRON             |
| 6767795 | 2004          | 15      | Renewing     | 1           | 6.70  | Highly reliable amorphous high-k gate dielectric ZrOXNY  | MICRON             |
| 6921702 | 2005          | 15      | Renewing     | 1           | 5.47  | Atomic layer deposited nanolaminates of HfO2/ZrO2 films as gate dielectrics  | MICRON             |
| 6660657 | 2003          | 15      | Renewing     | 1           | 0.68  | Methods of incorporating nitrogen into silicon-oxide-containing layers   | MICRON             |
| 6429061 | 2002          | 16      | Renewing     | 2           | 5.66  | Complimentary metal oxide semiconductor (cmos); producing higher performance device; forming a relaxed silicon germanium layer with isolation and well implant regions | IBM                |
| 6291845 | 2001          | 16      | Renewing     | 2           | 1.87  | Fully-dielectric-isolated FET technology   | STMICROELECTRONICS |
| 6841457 | 2005          | 16      | Renewing     | 2           | 1.42  | Use of hydrogen implantation to improve material properties of silicon-germanium-on-insulator material made by thermal diffusion                                       | IBM                |
| 6724008 | 2004          | 16      | Renewing     | 2           | 1.07  | Relaxed silicon germanium platform for high speed CMOS electronics and high speed analog circuits  | AMBERWAVE          |
| 6713326 | 2004          | 16      | Renewing     | 2           | 0.91  | Process for producing semiconductor article using graded epitaxial growth  | MIT                |
| 6524920 | 2003          | 17      | Exhausting   | 2           | 22.14 | Low temperature process for a transistor with elevated source and drain  | AMD INC            |
| 6300201 | 2001          | 17      | Exhausting   | 2           | 6.20  | Method to form a high K dielectric gate insulator layer, a metal gate structure, and self-aligned channel regions, post source/drain formation                         | CHARTERED          |
| 6194748 | 2001          | 17      | Exhausting   | 2           | 4.46  | MOSFET with suppressed gate-edge fringing field effect   | AMD INC            |
| 6171910 | 2001          | 17      | Exhausting   | 2           | 2.81  | Method for forming a semiconductor device  | MOTOROLA           |
| 6380043 | 2002          | 17      | Exhausting   | 2           | 1.71  | Low temperature process to form elevated drain and source of a field effect transistor having high-K gate dielectric   | AMD INC            |
| 6933525 | 2005          | 18      | Breakthrough | 2           | 1.12  | Display device and manufacturing method of the same  | HITACHI            |
| 7084428 | 2006          | 18      | Breakthrough | 2           | 0.88  | Transistor, integrated circuit, electro-optic device, electronic instrument and method of manufacturing a transistor   | SEIKO EPSON CORP   |
| 6218219 | 2001          | 18      | Breakthrough | 2           | 0.60  | Semiconductor device and fabrication method thereof  | S.E.L              |
| 6407431 | 2002          | 18      | Breakthrough | 2           | 0.48  | Semiconductor device and fabrication method thereof  | S.E.L              |
| 6762468 | 2004          | 18      | Breakthrough | 2           | 0.21  | Semiconductor device and method of manufacturing the same  | TOSHIBA            |
| 6251738 | 2001          | 19      | Exhausting   | 2           | 0.92  | Process for forming a silicon-germanium base of heterojunction bipolar transistor  | IBM                |
| 6521502 | 2003          | 20      | Renewing     | 2           | 12.74 | Solid phase epitaxy activation process for source/drain junction extensions and halo regions   | AMD INC            |
| 6365476 | 2002          | 20      | Renewing     | 2           | 8.07  | Laser thermal process for fabricating field-effect transistors   | ULTRATECH STEPPER  |

| Patent  | grant<br>year | Cluster | LC stage   | NMP<br>comp | PathC | Title  | Assignee                    |
|---------|---------------|---------|------------|-------------|-------|--|-----------------------------|
| 6660605 | 2003          | 20      | Renewing   | 2           | 6.72  | Method to fabricate optimal HDD with dual diffusion process to optimize transistor drive current junction capacitance, tunneling current and channel dopant loss   | TEXAS INSTRUMENTS           |
| 6225173 | 2001          | 20      | Renewing   | 2           | 5.80  | Recessed channel structure for manufacturing shallow source/drain extensions   | AMD INC                     |
| 6218250 | 2001          | 20      | Renewing   | 2           | 5.41  | Method and apparatus for minimizing parasitic resistance of semiconductor devices  | AMD INC                     |
| 6440793 | 2002          | 21      | Disruptive | 2           | 6.47  | Vertical MOSFET  | IBM                         |
| 6261894 | 2001          | 21      | Disruptive | 2           | 3.63  | Method for forming dual workfunction high-performance support MOSFETs in EDRAM arrays  | IBM                         |
| 6964897 | 2005          | 21      | Disruptive | 2           | 2.28  | SOI trench capacitor cell incorporating a low-leakage floating body array transistor   | IBM                         |
| 7122840 | 2006          | 21      | Disruptive | 2           | 1.23  | Image sensor with optical guard ring and fabrication method thereof  | TSMC                        |
| 7098146 | 2006          | 21      | Disruptive | 2           | 1.11  | Semiconductor device having patterned SOI structure and method for fabricating the same  | TOSHIBA                     |
| 6703648 | 2004          | 22      | Disruptive | 2           | 18.24 | Strained silicon PMOS having silicon germanium source/drain extensions and method for its fabrication  | AMD INC                     |
| 6743684 | 2004          | 22      | Disruptive | 2           | 14.29 | Method to produce localized halo for MOS transistor  | TEXAS INSTRUMENTS           |
| 6881632 | 2005          | 22      | Disruptive | 2           | 10.56 | Method of fabricating CMOS inverter and integrated circuits utilizing strained surface channel MOSFETS   | AMBERWAVE                   |
| 7074623 | 2006          | 22      | Disruptive | 2           | 8.70  | Methods of forming strained-semiconductor-on-insulator finFET device structures  | AMBERWAVE                   |
| 7122449 | 2006          | 22      | Disruptive | 2           | 7.24  | Methods of fabricating semiconductor structures having epitaxially grown source and drain elements   | AMBERWAVE                   |
| 6190977 | 2001          | 24      | Exhausting | 2           | 28.94 | Method for forming MOSFET with an elevated source/drain  | TEXAS INSTRUMENTS -<br>ACER |
| 6303450 | 2001          | 24      | Exhausting | 2           | 8.48  | CMOS device structures and method of making same   | IBM                         |
| 6284657 | 2001          | 25      | Mature     | 2           | 1.52  | Non-metallic barrier formation for copper damascene type interconnects   | CHARTERED                   |
| 7122442 | 2006          | 25      | Mature     | 2           | 0.47  | Method and system for dopant containment   | TEXAS INSTRUMENTS           |
| 6611045 | 2003          | 25      | Mature     | 2           | 0.17  | Method of forming an integrated circuit device using dummy features and structure thereof  | MOTOROLA                    |
| 6642579 | 2003          | 25      | Mature     | 2           | 0.16  | Method of reducing the extrinsic body resistance in a silicon-on-insulator body contacted MOSFET   | IBM                         |
| 6864155 | 2005          | 25      | Mature     | 2           | 0.14  | Methods of forming silicon-on-insulator comprising integrated circuitry, and wafer bonding methods of forming silicon-on-insulator comprising integrated circuitry | MICRON                      |
| 6555839 | 2003          | 26      | Renewing   | 2           | 1.43  | Buried channel strained silicon FET using a supply layer created through ion implantation  | AMBERWAVE                   |
| 6350993 | 2002          | 26      | Renewing   | 2           | 0.41  | High speed composite p-channel Si/SiGe heterostructure for field effect devices  | IBM                         |
| 6207977 | 2001          | 26      | Renewing   | 2           | 0.04  | Vertical MISFET devices  | IMEC                        |
| 6204126 | 2001          | 27      | Exhausting | 2           | 2.52  | Method to fabricate a new structure with multi-self-aligned for split-gate flash   | TSMC                        |
| 6573126 | 2003          | 28      | Renewing   | 2           | 0.81  | Process for producing semiconductor article using graded epitaxial growth  | MIT                         |

| Patent  | grant<br>year | Cluster | LC stage   | NMP<br>comp | PathC | Title  | Assignee           |
|---------|---------------|---------|------------|-------------|-------|--|--------------------|
| 6323108 | 2001          | 28      | Renewing   | 2           | 0.24  | Fabrication ultra-thin bonded semiconductor layers   | US NAVY            |
| 6261929 | 2001          | 28      | Renewing   | 2           | 0.23  | Methods of forming a plurality of semiconductor layers using spaced trench arrays  | NORTHCAR. ST. UNI. |
| 6191007 | 2001          | 28      | Renewing   | 2           | 0.12  | Method for manufacturing a semiconductor substrate   | DENSO CORP LTD     |
| 6235567 | 2001          | 28      | Renewing   | 2           | 0.06  | Silicon-germanium bicmos on soi  | IBM                |
| 6413802 | 2002          | 29      | Disruptive | 2           | 29.61 | Finfet transistor structures having a double gate channel extending vertically from a substrate and methods of manufacture | UNIV OF CALIFORNIA |
| 6214670 | 2001          | 29      | Disruptive | 2           | 18.48 | Method for manufacturing short-channel, metal-gate CMOS devices with superior hot carrier performance                      | TSMC               |
| 6686231 | 2004          | 29      | Disruptive | 2           | 13.00 | Damascene gate process with sacrificial oxide in semiconductor devices   | AMD INC            |
| 7084018 | 2006          | 29      | Disruptive | 2           | 10.55 | Sacrificial oxide for minimizing box undercut in damascene FinFET  | AMD INC            |
| 6962843 | 2005          | 29      | Disruptive | 2           | 8.79  | Method of fabricating a finfet   | IBM                |

#### A.4.6 SRTA tables at the firm level

In this section we report the SRTA values calculated for a selection of firms from the US, Japan, Korea, Taiwan and Singapore. To keep the analysis short we do that only for the last three periods. Tables from Table 31 to Table 38 reports the SRTA values for the main US, Japanese, Taiwanese, Korean and Singaporean players over time. We highlight values of the SRTA greater than 0.2 in bold. Firms are distinguished between new and incumbent innovators and also based on their business area (IDM=Integrated Device Manufacturer, GRO=Government Research Organization, NGRO=Non-Governmental Research Organization, Equipm.=Equipment supplier). The tables confirm comparative technological advantage patterns as discussed in Section 4.5.2. However, they provide further details for those interested to track specialization trends for particular firms or research institutes.

**Table 31: SRTA for the top Taiwanese, Korean and Singaporean firms (1991-1995)**

| Company            | New Inn vs Inc | Type    | #Patents | Disruptive | Early growth | Mature | Renewing | Exhausting |
|--------------------|----------------|---------|----------|------------|--------------|--------|----------|------------|
| UMC (TW)           | New innovator  | Foundry | 31       | -0,477     | 0,230        | -0,087 | 0,597    | 0,597      |
| SAMSUNG (KR)       | Incumbent      | IDM     | 8        | 0,046      | 0,300        | -0,365 | -1,000   | -1,000     |
| TITRI (TW)         | Incumbent      | GRO     | 7        | 0,112      | -0,171       | -0,306 | 0,490    | -1,000     |
| HYUNDAI ELEC. (KR) | New innovator  | IDM     | 7        | -0,523     | 0,359        | -0,306 | 0,708    | -1,000     |
| LG ELEC. (KR)      | New innovator  | IDM     | 7        | -0,230     | -0,171       | 0,360  | -1,000   | -1,000     |
| TSMC (TW)          | New innovator  | Foundry | 6        | -0,155     | 0,245        | 0,107  | -1,000   | -1,000     |
| CHARTERED (SG)     | New innovator  | Foundry | 4        | -1,000     | 0,664        | -1,000 | -1,000   | -1,000     |
| KETRI (KR)         | Incumbent      | GRO     | 3        | 0,188      | -1,000       | 0,107  | -1,000   | -1,000     |
| WINBOND (TW)       | New innovator  | IDM     | 2        | -1,000     | -1,000       | 0,576  | -1,000   | -1,000     |

**Table 32: SRTA for the top US and Japanese players (1991-1995)**

| Company                 | New Inn vs Inc | Type    | #Patents | Disruptive | Early growth | Mature | Renewing | Exhausting |
|-------------------------|----------------|---------|----------|------------|--------------|--------|----------|------------|
| TEXAS INSTR. (US)       | Incumbent      | IDM     | 39       | -0,053     | -0,223       | 0,177  | -0,312   | 0,355      |
| MOTOROLA (US)           | Incumbent      | IDM     | 38       | -0,040     | -0,211       | -0,010 | 0,235    | 0,623      |
| MICRON (US)             | New innovator  | IDM     | 38       | 0,096      | 0,132        | -0,546 | 0,235    | 0,037      |
| IBM (US)                | Incumbent      | IDM     | 35       | 0,159      | -0,005       | -0,221 | -1,000   | -1,000     |
| MITSUBISHI (JP)         | Incumbent      | IDM     | 33       | -0,073     | -0,052       | 0,189  | -0,234   | -1,000     |
| TOSHIBA (JP)            | Incumbent      | IDM     | 33       | -0,202     | -0,538       | 0,340  | 0,301    | -1,000     |
| NEC (JP)                | Incumbent      | IDM     | 22       | -0,335     | -0,052       | 0,375  | -1,000   | -1,000     |
| AT&T (US)               | Incumbent      | IDM     | 17       | 0,016      | 0,186        | -0,207 | 0,093    | -1,000     |
| SONY CORP (JP)          | Incumbent      | IDM     | 17       | -0,051     | -0,264       | 0,273  | -1,000   | -1,000     |
| FUJITSU (JP)            | Incumbent      | IDM     | 13       | 0,083      | -0,135       | -0,076 | 0,223    | -1,000     |
| HITACHI (JP)            | Incumbent      | Equipm. | 11       | 0,089      | -0,052       | 0,007  | -1,000   | -1,000     |
| NATIONAL SEMICOND. (US) | Incumbent      | IDM     | 11       | -0,002     | -1,000       | 0,340  | -1,000   | -1,000     |
| HARRIS (US)             | Incumbent      | User    | 7        | 0,374      | -1,000       | -1,000 | -1,000   | -1,000     |
| LSI LOGIC (US)          | Incumbent      | Fabless | 7        | 0,305      | -0,171       | -1,000 | -1,000   | -1,000     |
| APPLIED MATERIALS (US)  | Incumbent      | Equipm. | 6        | 0,188      | 0,245        | -1,000 | -1,000   | -1,000     |
| HUGHES (US)             | Incumbent      | User    | 6        | -0,465     | -1,000       | 0,425  | 0,547    | -1,000     |
| MATSUSHITA (JP)         | Incumbent      | IDM     | 6        | -0,465     | 0,245        | 0,107  | -1,000   | 0,744      |

| Company               | New Inn vs Inc | Type | #Patents | Disruptive | Early growth | Mature | Renewing | Exhausting |
|-----------------------|----------------|------|----------|------------|--------------|--------|----------|------------|
| OKI ELECTRIC (JP)     | Incumbent      | IDM  | 6        | -1,000     | 0,245        | 0,425  | -1,000   | -1,000     |
| SHARP (JP)            | Incumbent      | IDM  | 6        | 0,046      | -0,096       | -0,234 | 0,547    | -1,000     |
| SIEMENS (DE)          | Incumbent      | IDM  | 6        | 0,046      | -0,096       | 0,107  | -1,000   | -1,000     |
| HONEYWELL (US)        | Incumbent      | IDM  | 5        | 0,274      | -0,005       | -1,000 | -1,000   | -1,000     |
| SEIKO EPSON (JP)      | Incumbent      | IDM  | 5        | -0,390     | 0,597        | -1,000 | -1,000   | -1,000     |
| SEMICOND. ENERGY (JP) | Incumbent      | NGRO | 5        | -0,065     | 0,329        | -0,147 | -1,000   | -1,000     |

**Table 33: SRTA for the top Taiwanese, Korean and Singaporean players (1996-2000)**

| Company            | New Inn vs Inc | Type    | #Patents | Break-through | Disruptive | Early growth | Mature | Renewing | Exhausting |
|--------------------|----------------|---------|----------|---------------|------------|--------------|--------|----------|------------|
| TSMC (TW)          | Incumbent      | Foundry | 92       | -0,429        | -0,310     | -0,018       | 0,004  | 0,361    | -1,000     |
| UMC (TW)           | Incumbent      | Foundry | 77       | -0,725        | -0,653     | 0,089        | -0,248 | 0,101    | 0,748      |
| SAMSUNG (KR)       | Incumbent      | IDM     | 31       | -0,117        | -1,000     | 0,029        | 0,636  | -0,033   | -1,000     |
| CHARTERED (SG)     | Incumbent      | Foundry | 29       | -0,224        | -0,284     | -0,151       | 0,231  | 0,385    | 0,804      |
| VANGUARD (TW)      | New innovator  | Foundry | 25       | -1,000        | -1,000     | 0,160        | -1,000 | 0,075    | -1,000     |
| LG ELEC. (KR)      | Incumbent      | IDM     | 21       | 0,187         | -0,130     | -0,122       | 0,377  | 0,161    | -1,000     |
| HYUNDAI ELEC. (KR) | Incumbent      | IDM     | 17       | -0,470        | -1,000     | 0,184        | -1,000 | -0,402   | -1,000     |
| ACER (TW)          | New innovator  | IDM     | 13       | -1,000        | 0,426      | 0,065        | -1,000 | 0,055    | -1,000     |
| TITRI (TW)         | Incumbent      | GRO     | 9        | -0,190        | -1,000     | -0,045       | -1,000 | 0,415    | -1,000     |
| MOSEL VITELIC (TW) | New innovator  | IDM     | 6        | 0,011         | -1,000     | -0,098       | -1,000 | 0,415    | -1,000     |
| WINBOND (TW)       | Incumbent      | IDM     | 5        | 0,102         | -1,000     | -0,007       | -1,000 | 0,184    | -1,000     |

**Table 34: SRTA for the top US and Japanese players (1996-2000)**

| Company                 | New Inn vs Inc | Type    | #Patents | Break-through | Disruptive | Early growth | Mature | Renewing | Exhausting |
|-------------------------|----------------|---------|----------|---------------|------------|--------------|--------|----------|------------|
| AMD (US)                | Incumbent      | IDM     | 93       | -0,117        | -0,704     | 0,029        | 0,332  | 0,111    | -1,000     |
| MICRON (US)             | Incumbent      | IDM     | 66       | 0,011         | -0,606     | 0,068        | 0,169  | -0,205   | -1,000     |
| NEC (JP)                | Incumbent      | IDM     | 49       | 0,239         | -0,504     | -0,031       | -0,027 | -0,059   | -1,000     |
| IBM (US)                | Incumbent      | IDM     | 37       | 0,140         | -0,392     | 0,053        | 0,113  | -0,436   | -1,000     |
| TEXAS INSTR. (US)       | Incumbent      | IDM     | 36       | -0,190        | -1,000     | 0,086        | 0,126  | 0,004    | -1,000     |
| MOTOROLA (US)           | Incumbent      | IDM     | 35       | 0,102         | -1,000     | -0,007       | 0,598  | -0,093   | -1,000     |
| TOSHIBA (JP)            | Incumbent      | IDM     | 25       | -0,010        | -1,000     | 0,084        | -1,000 | -0,069   | -1,000     |
| MITSUBISHI (JP)         | Incumbent      | IDM     | 21       | 0,078         | -1,000     | 0,046        | -1,000 | 0,018    | -1,000     |
| MATSUSHITA (JP)         | Incumbent      | IDM     | 18       | 0,463         | -1,000     | -0,098       | 0,441  | -1,000   | -1,000     |
| NATIONAL SEMICOND. (US) | Incumbent      | IDM     | 17       | 0,433         | -1,000     | -0,128       | -1,000 | -0,079   | -1,000     |
| LSI LOGIC (US)          | Incumbent      | Fabless | 16       | -1,000        | -1,000     | 0,244        | -1,000 | -1,000   | -1,000     |
| SHARP (JP)              | Incumbent      | IDM     | 15       | 0,241         | -1,000     | -0,007       | -1,000 | -0,016   | -1,000     |
| INTEL (US)              | Incumbent      | IDM     | 12       | 0,343         | -1,000     | -0,292       | -1,000 | 0,415    | -1,000     |
| LUCENT (US)             | New innovator  | User    | 12       | -1,000        | -1,000     | 0,202        | -1,000 | -0,246   | -1,000     |

| Company                 | New Inn vs Inc | Type       | #Patents | Break-through | Disruptive | Early growth | Mature | Renewing | Exhausting |
|-------------------------|----------------|------------|----------|---------------|------------|--------------|--------|----------|------------|
| SONY CORP (JP)          | Incumbent      | IDM        | 11       | 0,252         | -1,000     | 0,023        | 0,617  | -1,000   | -1,000     |
| HITACHI (JP)            | Incumbent      | Equipm.    | 10       | -0,240        | 0,236      | 0,070        | -1,000 | -0,159   | -1,000     |
| VLSI TECH (US)          | Incumbent      | IDM        | 9        | 0,154         | -1,000     | 0,046        | -1,000 | -0,107   | -1,000     |
| SEMICOND. ENERGY (JP)   | Incumbent      | NGRO       | 7        | -1,000        | -1,000     | 0,244        | -1,000 | -1,000   | -1,000     |
| YAMAHA (JP)             | Incumbent      | IDM        | 7        | -0,066        | -1,000     | -0,031       | -1,000 | 0,349    | -1,000     |
| SIEMENS (DE)            | Incumbent      | IDM        | 6        | 0,508         | 0,459      | -0,570       | 0,771  | -1,000   | -1,000     |
| APPLIED MATERIALS (US)  | Incumbent      | Equipm.    | 5        | 0,421         | -1,000     | -0,007       | -1,000 | -1,000   | -1,000     |
| UNIV CALIFORNIA (US)    | Incumbent      | University | 5        | -1,000        | 0,528      | -0,007       | -1,000 | 0,184    | -1,000     |
| SANYO ELECTRIC (JP)     | Incumbent      | IDM        | 5        | 0,102         | 0,732      | -0,207       | -1,000 | -1,000   | -1,000     |
| AMERICAN SUPERCOND.(US) | New innovator  | User       | 5        | -1,000        | 0,883      | -1,000       | -1,000 | -1,000   | -1,000     |
| FOVEONICS (US)          | New innovator  | User       | 5        | -1,000        | -1,000     | 0,244        | -1,000 | -1,000   | -1,000     |

**Table 35: SRTA for the top Taiwanese, Korean and Singaporean players (2001-2006 - Main component of the network of main paths)**

| Company            | New Inn vs Inc | Type    | #Patents | Disruptive | Mature | Renewing | Exhausting |
|--------------------|----------------|---------|----------|------------|--------|----------|------------|
| TSMC (TW)          | Incumbent      | Foundry | 13       | 0,196      | 0,397  | -0,165   | -1,000     |
| SAMSUNG (KR)       | Incumbent      | IDM     | 9        | -1,000     | 0,540  | 0,095    | -1,000     |
| CHARTERED (SG)     | Incumbent      | Foundry | 4        | -0,017     | 0,580  | -0,125   | -1,000     |
| UMC (TW)           | Incumbent      | Foundry | 4        | 0,318      | -1,000 | -0,125   | -1,000     |
| HYUNDAI ELEC. (KR) | Incumbent      | IDM     | 3        | -1,000     | -1,000 | 0,217    | -1,000     |
| VANGUARD (TW)      | Incumbent      | Foundry | 1        | 0,589      | -1,000 | -1,000   | -1,000     |
| HYNIX (KR)         | New innovator  | IDM     | 1        | -1,000     | -1,000 | 0,217    | -1,000     |

**Table 36: SRTA for the top US and Japanese players (2001-2006 - Main component of the network of main paths)**

| Company                | New Inn vs Inc | Type    | #Patents | Disruptive | Mature | Renewing | Exhausting |
|------------------------|----------------|---------|----------|------------|--------|----------|------------|
| MICRON (US)            | Incumbent      | IDM     | 75       | -0,276     | -0,666 | 0,133    | -1,000     |
| AMD (US)               | Incumbent      | IDM     | 31       | -0,068     | 0,489  | -0,141   | 0,509      |
| IBM (US)               | Incumbent      | IDM     | 22       | -0,175     | 0,345  | 0,029    | -1,000     |
| APPLIED MATERIALS (US) | Incumbent      | Equipm. | 17       | 0,494      | -1,000 | -0,691   | 0,578      |
| TEXAS INSTR. (US)      | Incumbent      | IDM     | 15       | 0,015      | -1,000 | 0,065    | -1,000     |
| MOTOROLA (US)          | Incumbent      | IDM     | 14       | -0,567     | 0,036  | 0,142    | -1,000     |
| SHARP (JP)             | Incumbent      | IDM     | 11       | -1,000     | -1,000 | -0,005   | 0,841      |
| INFINEON (DE)          | Incumbent      | IDM     | 4        | -0,017     | -1,000 | 0,077    | -1,000     |
| NOVELIUS SYSTEMS (US)  | New innovator  | Equipm. | 4        | 0,487      | -1,000 | -0,440   | -1,000     |
| LAM (US)               | Incumbent      | Equipm. | 3        | 0,589      | -1,000 | -1,000   | -1,000     |
| MATSUSHITA (JP)        | Incumbent      | IDM     | 3        | 0,126      | -1,000 | 0,018    | -1,000     |
| GENUS (US)             | New innovator  | Equipm. | 3        | -1,000     | -1,000 | 0,217    | -1,000     |

**Table 37: SRTA for the top Taiwanese, Korean and Singaporean players  
(2001-2006 - Second component of the network of main paths)**

| Company            | New Inn vs Inc | Type    | #Patents | Breakthrough | Disruptive | Mature | Renewing | Exhausting |
|--------------------|----------------|---------|----------|--------------|------------|--------|----------|------------|
| TSMC (TW)          | Incumbent      | Foundry | 40       | -0,700       | 0,100      | -0,131 | 0,000    | 0,084      |
| SAMSUNG (KR)       | Incumbent      | IDM     | 18       | 0,594        | -0,347     | -1,000 | -0,091   | -1,000     |
| LG PHILIPS (KR)    | New innovator  | IDM     | 13       | 0,752        | -1,000     | -1,000 | -1,000   | -1,000     |
| UMC (TW)           | Incumbent      | Foundry | 10       | -1,000       | 0,024      | -1,000 | 0,333    | -1,000     |
| HYUNDAI ELEC. (KR) | Incumbent      | IDM     | 9        | -1,000       | 0,152      | 0,262  | -0,286   | -1,000     |
| CHARTERED (SG)     | Incumbent      | Foundry | 7        | -1,000       | -0,001     | 0,374  | -0,167   | 0,742      |
| HANN STAR (TW)     | New innovator  | IDM     | 5        | 0,752        | -1,000     | -1,000 | -1,000   | -1,000     |
| KETRI (KR)         | Incumbent      | GRO     | 3        | -1,000       | 0,076      | -1,000 | 0,250    | -1,000     |
| MACRONIOX (TW)     | Incumbent      | IDM     | 3        | -1,000       | 0,076      | -1,000 | 0,250    | -1,000     |
| CHUNGHWA (TW)      | New innovator  | IDM     | 3        | 0,752        | -1,000     | -1,000 | -1,000   | -1,000     |
| HYNIX (KR)         | New innovator  | IDM     | 3        | -1,000       | -0,264     | 0,673  | 0,250    | -1,000     |
| TITRI (TW)         | Incumbent      | GRO     | 2        | 0,559        | -1,000     | -1,000 | -1,000   | 0,919      |
| VANGUARD (TW)      | Incumbent      | Foundry | 2        | -1,000       | 0,272      | -1,000 | -1,000   | -1,000     |
| AU OPTRONIC (TW)   | New innovator  | IDM     | 2        | 0,752        | -1,000     | -1,000 | -1,000   | -1,000     |

**Table 38: SRTA for the top US and Japanese players  
(2001-2006 - Second component of the network of main paths)**

| Company                | New Inn vs Inc | Type       | #Patents | Breakthrough | Disruptive | Mature | Renewing | Exhausting |
|------------------------|----------------|------------|----------|--------------|------------|--------|----------|------------|
| AMD (US)               | Incumbent      | IDM        | 81       | -1,000       | 0,038      | -0,026 | 0,152    | 0,401      |
| IBM (US)               | Incumbent      | IDM        | 73       | -0,348       | 0,118      | -0,226 | -0,187   | 0,129      |
| TOSHIBA (JP)           | Incumbent      | IDM        | 33       | -0,218       | 0,003      | -1,000 | 0,250    | -1,000     |
| TEXAS INSTR. (US)      | Incumbent      | IDM        | 23       | -1,000       | -0,046     | 0,335  | 0,270    | -1,000     |
| SEMICOND. ENERGY (JP)  | Incumbent      | NGRO       | 18       | 0,725        | -0,823     | -1,000 | -0,565   | -1,000     |
| MICRON (US)            | Incumbent      | IDM        | 17       | -1,000       | 0,061      | -0,050 | 0,190    | -1,000     |
| NEC (JP)               | Incumbent      | IDM        | 13       | -0,296       | 0,193      | 0,084  | -1,000   | -1,000     |
| AMBERWAVE SYSTEMS (US) | New innovator  | Equipm.    | 13       | -1,000       | 0,147      | -1,000 | 0,071    | -1,000     |
| INTEL (US)             | Incumbent      | IDM        | 12       | -1,000       | 0,186      | -1,000 | -0,412   | 0,595      |
| MITSUBISHI (JP)        | Incumbent      | IDM        | 9        | 0,404        | -0,264     | 0,547  | -0,286   | -1,000     |
| SHARP (JP)             | Incumbent      | IDM        | 9        | -0,120       | -0,675     | -1,000 | 0,591    | -1,000     |
| MATSUSHITA (JP)        | Incumbent      | IDM        | 7        | -1,000       | 0,272      | -1,000 | -1,000   | -1,000     |
| FUJITSU (JP)           | Incumbent      | IDM        | 6        | 0,082        | -0,264     | 0,673  | -0,091   | -1,000     |
| LSI LOGIC (US)         | Incumbent      | Fabless    | 6        | -1,000       | 0,076      | 0,673  | -1,000   | -1,000     |
| MIT (US)               | Incumbent      | University | 6        | -1,000       | -0,067     | -1,000 | 0,429    | -1,000     |
| CANON (JP)             | Incumbent      | User       | 5        | -1,000       | 0,272      | -1,000 | -1,000   | -1,000     |
| HITACHI (JP)           | Incumbent      | Equipm.    | 5        | 0,171        | 0,024      | -1,000 | 0,000    | -1,000     |
| HUGHES (US)            | Incumbent      | User       | 5        | -1,000       | 0,272      | -1,000 | -1,000   | -1,000     |
| MOTOROLA (US)          | Incumbent      | IDM        | 5        | -1,000       | 0,024      | 0,509  | -1,000   | 0,809      |
| FREESCALE (US)         | New innovator  | IDM        | 5        | -1,000       | 0,166      | -1,000 | 0,000    | -1,000     |
| INFINEON (DE)          | Incumbent      | IDM        | 4        | -1,000       | 0,134      | 0,587  | -1,000   | -1,000     |
| APPLIED MATERIALS (US) | Incumbent      | Equipm.    | 3        | 0,650        | -0,264     | -1,000 | -1,000   | -1,000     |
| OKI ELECTRIC (JP)      | Incumbent      | IDM        | 3        | -1,000       | -0,264     | -1,000 | 0,538    | -1,000     |
| SONY CORP (JP)         | Incumbent      | IDM        | 3        | -1,000       | 0,076      | -1,000 | 0,250    | -1,000     |
| AGERE SYSTEM (US)      | New innovator  | Fabless    | 3        | -1,000       | 0,076      | -1,000 | 0,250    | -1,000     |
| E INK (US)             | New innovator  | IDM        | 3        | 0,752        | -1,000     | -1,000 | -1,000   | -1,000     |

|                      |               |      |   |        |        |        |        |        |
|----------------------|---------------|------|---|--------|--------|--------|--------|--------|
| HONEYWELL (US)       | Incumbent     | User | 3 | -1,000 | 0,272  | -1,000 | -1,000 | -1,000 |
| RENESAS ELECTR. (JP) | New innovator | IDM  | 3 | -1,000 | -0,264 | 0,673  | 0,250  | -1,000 |



## APPENDIX TO CHAPTER 5

### A.5.1 Ranking of observations by knowledge depth

Table 39: top 10% observations by ln(KD)

| Time period | Company label | Country | Business cat. | ln(KD)    | KD        | KB        | PC |
|-------------|---------------|---------|---------------|-----------|-----------|-----------|----|
| 1995        | MOTOROLA      | US      | IDM           | -2.467234 | 0.0848191 | 0.9521546 | 45 |
| 1990        | TEXASINST     | US      | IDM           | -2.827807 | 0.0591424 | 0.9250826 | 56 |
| 1995        | ANALOGDEV     | US      | IDM           | -2.942266 | 0.052746  | 0         | 1  |
| 1995        | FUJITSU       | JP      | IDM           | -2.958791 | 0.0518816 | 0.8963358 | 14 |
| 1995        | AMD           | US      | IDM           | -3.080825 | 0.0459213 | 0.7635733 | 6  |
| 1995        | SPRAGUE       | US      | IDM           | -3.116289 | 0.0443213 | 0         | 1  |
| 1995        | TEXASINST     | US      | IDM           | -3.262161 | 0.0383055 | 0.9281647 | 43 |
| 2000        | AMD           | US      | IDM           | -3.466486 | 0.0312266 | 0.905466  | 93 |
| 2000        | UMC           | TW      | Foundry       | -3.493906 | 0.030382  | 0.8842815 | 77 |
| 1985        | EATON         | US      | User          | -3.521066 | 0.0295679 | 0.2231109 | 4  |
| 1990        | TOSHIBA       | JP      | IDM           | -3.535164 | 0.029154  | 0.9613467 | 74 |
| 1995        | MITSUBISHI    | JP      | IDM           | -3.635369 | 0.0263742 | 0.8210678 | 33 |
| 1995        | KE&T          | KR      | Res.prov.     | -3.748003 | 0.0235648 | 0.6802697 | 4  |
| 1985        | IBM           | US      | IDM           | -3.760301 | 0.0232767 | 0.8622844 | 96 |
| 1995        | MICRON        | US      | IDM           | -3.983842 | 0.018614  | 0.8571347 | 41 |
| 1995        | HYUNDAI       | KR      | IDM           | -4.020535 | 0.0179434 | 0.8265074 | 8  |
| 1995        | UMC           | TW      | Foundry       | -4.050233 | 0.0174183 | 0.8781988 | 32 |
| 2000        | MICRON        | US      | IDM           | -4.120855 | 0.0162306 | 0.9006752 | 66 |
| 1990        | MOTOROLA      | US      | IDM           | -4.195992 | 0.0150558 | 0.9330899 | 35 |
| 1985        | RCA           | US      | IDM           | -4.212543 | 0.0148087 | 0.8940903 | 28 |
| 1995        | TSMC          | TW      | Foundry       | -4.299595 | 0.0135741 | 0.7682303 | 8  |
| 2000        | TSMC          | TW      | Foundry       | -4.333607 | 0.0131201 | 0.9048561 | 88 |
| 1995        | LG            | KR      | IDM           | -4.33592  | 0.0130898 | 0.8020592 | 9  |
| 2000        | CHARTERED     | SG      | Foundry       | -4.474548 | 0.0113954 | 0.9203129 | 29 |
| 1985        | AT&T          | US      | IDM           | -4.487476 | 0.011249  | 0.8634174 | 42 |
| 1995        | IBM           | US      | IDM           | -4.487646 | 0.0112471 | 0.8455559 | 36 |
| 1985        | TOSHIBA       | JP      | IDM           | -4.52588  | 0.0108252 | 0.8843337 | 57 |
| 1995        | NEC           | JP      | IDM           | -4.596872 | 0.0100833 | 0.784089  | 23 |
| 1985        | FUJI_XEROX    | JP      | User          | -4.653471 | 0.0095285 | 0         | 1  |
| 1995        | LSILOGIC      | US      | Fabless       | -4.697524 | 0.0091178 | 0.6320425 | 7  |
| 1995        | TOSHIBA       | JP      | IDM           | -4.763703 | 0.0085339 | 0.8766038 | 35 |
| 1990        | HITACHI       | JP      | Supplier      | -4.765155 | 0.0085216 | 0.9206975 | 43 |
| 1990        | NCR           | US      | User          | -4.809824 | 0.0081493 | 0.620767  | 9  |
| 1985        | HITACHI       | JP      | Supplier      | -4.83981  | 0.0079086 | 0.9408442 | 24 |

## A.5.2 Empirical conditional probabilities of survival

To provide preliminary insights about the strength and functional form of the association between the probability of survival and our explanatory variables we computed the empirical conditional probabilities of survival for each of our variables. We broke continuous variables (KB, KD, PC) into bins and calculated the survival rate (i.e. the ratio between the number of survivors and the number of observations) within each bin. The conditional probability of survival for categorical variables is straightforward and computed as the survival rate within the given category. Figure 55 shows the results.

The conditional probability of survival for knowledge breadth follows a very irregular trend up to the fifth bin (i.e. up to  $KB < 0.4$ ). This might be an interesting insight in itself or, contrarily it might be purely induced by the scarcity of observations for values of KB in the range  $0 \leq KB < 0.4$  (i.e. bins 2-4). Indeed from the fifth bin onwards  $P(S|KB)$  follows an upward sloping trend, with the exception of bin 9, in which it mildly decreases. The irregular trend poses some challenges in terms of the estimation of the functional form of the relationship between diversification and survival, particularly because we cannot safely rely on the information provided within the 0-0.4 range of KB.

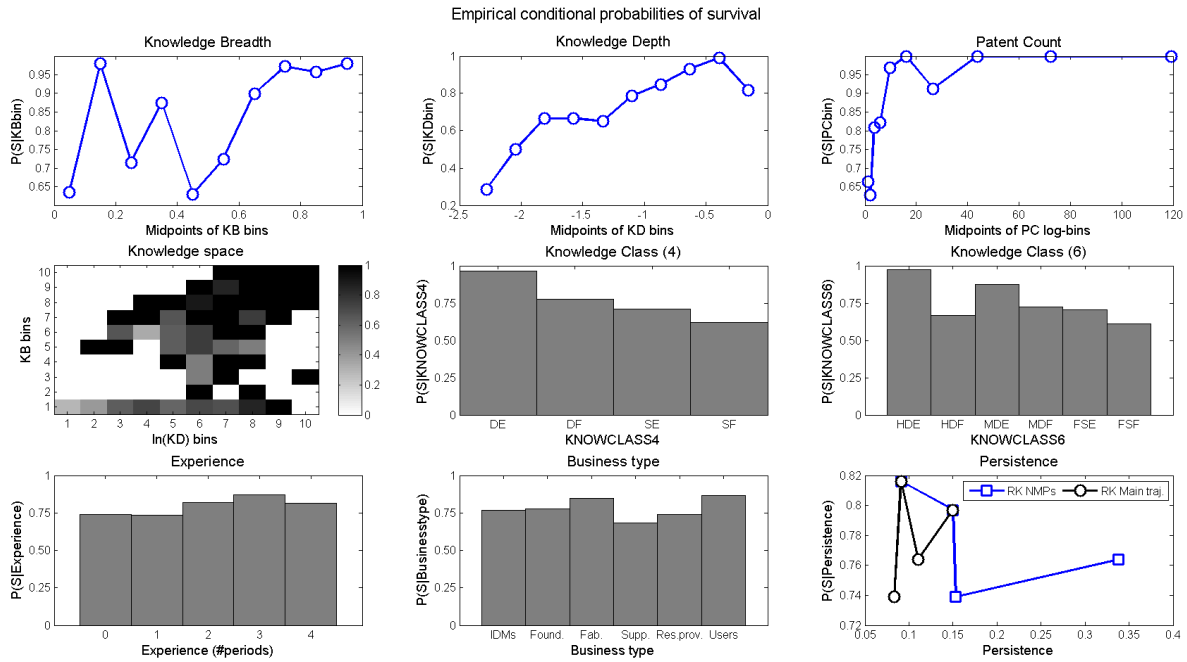


Figure 55: Empirical conditional probabilities of survival

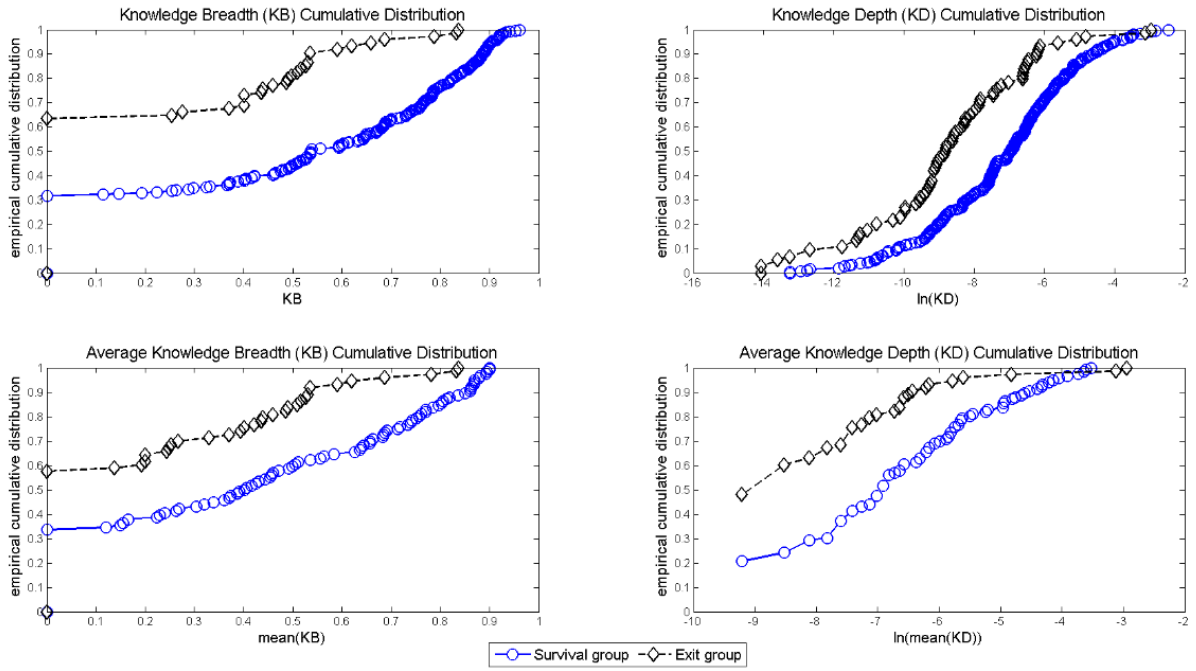
The relationship between survival and knowledge accumulation seems to be clearer. However also in this case some of the bins have too little observations to provide sound conjectures. These are bins 1, 2 and 10, in which there are only 7, 6 and 11 observations respectively. If we do not consider those bins, the observed probability of survival conditional to the level of knowledge accumulation seems to show a similar dynamic as for the knowledge breadth, with a clear increasing trend appearing only after a given threshold, which happens to be once again the fifth bin.

The size of the technological knowledge base seems to positively affect survival. In this case, given that the distribution of PC is highly skewed to the left we computed log-bins (i.e. bins of increasing width), in order to obtain more reliable and meaningful survival rates. The second

row of subplots of Figure 55 shows that larger values of both KB and KD seem to be associated with higher survival rates. Indeed the upper right corner of the KB-KD knowledge space has survival rates of 100% or close. Note that white spots in the figure indicate lack of observations in that region of the knowledge space. The business model followed and the years of experience does not seem to have any effect on survival chances. Furthermore, contrary to what expected a clear relationship between survival and knowledge persistence does not emerge if we only look at the empirical conditional probability observed in our sample.

### **A.5.3 Kolmogorov-Smirnof, Anderson-Darling and Student's T tests of the associations of knowledge breadth and depth with innovative survival**

Here we assess whether diversification and knowledge accumulation alone are associated with survival, without controlling for other explanations or determine the functional form of the relationship. This boils down to evaluate the correlation between the binary variable *Survival* and the continuous variables KB and KD. The simplest way to do it is to test whether the probability to observe larger values of KB or KD in the sample of innovators that survive is statistically larger than for the sample of innovators that exit. The top panels of Figure 56 plot the empirical cumulative distribution functions (ECDFs) of KB and the logged-KD for the two samples. The ECDF computes the observed probability of finding values of the given variable lower or equal than a given level. It does so by calculating the left-cumulated frequency of observations. For this purpose, we run two tests, the two-sample Kolmogorov-Smirnoff test and the Anderson-Darling test to assess whether the difference between the distributions of the two groups is statistically significant. The distribution of KB and KD for the survival group might be inflated by the values of the two variables for those players that survive for several periods. In order to assess whether this effect biases the results we first perform the tests for the distributions of KB and KD accounting for all observations and giving them equal weight. Then we do the same but this time computing a distribution in which each player (not each observation) has equal weight. In other words, we first calculate the average values of KB and KD for surviving players and then compute the ECDFs for the average values. Obviously the average values of KB and KD for players that survive for only one period is trivially equal to the values scored in that period. These ECDFs are plotted in the bottom panels of Figure 56.



**Figure 56: Empirical left-cumulative distribution functions of KB and KD for surviving and exiting core innovators**

In all four panels, we can see that, for the whole range of values the distribution of knowledge breadth and logged-depth for the survival group is stochastically dominated by the distribution of the exit group. This means that the probability to find larger values of KB and KD is higher for the survival group than for the exit group, for any value of KB and KD larger than zero. For instance, one can immediately see from the upper-left panel that larger values of knowledge breadth are more likely to be found in the survival group, where about 40% of the observations have a value of KB larger than 0.5, whereas this holds for only less than 20% of the observations in the exit group. Moreover, the crossing point with the vertical axis for the ECDF of KB shows that the probability to find fully specialized core innovators (i.e. with KB=0) is only about 35% in the survival group and nearly 60% in the exit one. This holds whether we filter the effect of long-lasting survival or not. We computed the probabilities to find observations with values of knowledge breadth and depth above the arithmetic and geometric mean<sup>26</sup> of the distribution respectively. They further confirm the visual impression that the distribution of knowledge depth is more skewed than the one of knowledge breadth. Fifty-three percent of all observations have a value higher than the arithmetic mean of KB and sixty percent higher than its geometric mean. To the contrary, only 19% of the observations have a value larger than the arithmetic mean of knowledge depth and 54% have a KD larger than the geometric mean. This means that for a randomly selected core innovator is easier to outperform the average competitor in terms of knowledge breadth than depth. We compute the same distributions for PC but not plot them here for the sake of synthesis. We perform three tests, the two-sample Kolmogorov-Smirnov test, the Anderson-Darling test and the t-test of paired differences, to compare whether the differences between the survival and exit group in terms of KB and KD are statistically significant. The

<sup>26</sup> The geometric mean is the mean of the log-transformed distribution

tests confirm that surviving innovators are statistically more likely to have broader and deeper knowledge base than exit ones. Table 40 summarizes the results for the two-sample Kolmogorov-Smirnoff test (KS-test). The null hypothesis of the KS-test is that the two groups belong to the same population, i.e. that there are no statistically significant differences between the ECDFs of the survival and exit group. The alternative hypothesis is that the ECDFs of the survival group is statistically smaller than the ECDF of the exit group (i.e. that the right tail of the distribution is heavier for the survival group than for the exit one).  $F_s(x)$  stands for the cumulative distribution function of the survival group, i.e. the share of observations in the survival group with a value lower or equal than  $x$ .  $F_e(x)$  is the cumulative distribution of the exit group. The t-statistic of the test corresponds to the highest difference found between the two curves. This is compared to the highest difference that could be found if one would randomly select two groups out of the same population. The value  $x^*$  corresponds to the value (of KB, KD, KB\*KD or PC) for which the largest possible difference  $F_e(x^*) - F_s(x^*)$  is found. The value  $s^*$  is equal to the ratio  $x^*/\max(x)$  and is helpful to assess how large  $x^*$  is. The KS-tests confirms that surviving core innovators are more likely to have large values of KB and KD than exiting ones.

We perform the same kind of analysis with the Anderson-Darling test (AD test). The KS test does not take into account at which level of the variable under consideration the highest difference in the two ECDFs is achieved. Since by definition the two curves have to converge to one, it is unlikely for  $x^*$  and  $s^*$  to have a large value. For our purposes, the AD test has the important advantage to give more weight to differences between the curves happening closer to the convergence point. The AD test results are reported in Table 40. They confirm the findings from the KS test.

Finally, we performed one last test based on a different approach. So far, we have compared probabilities to observe diversification and knowledge accumulation across surviving and exiting core innovators. However, there might be firm-specific effects that we are not capturing with the KS and AD tests. We therefore want to compare exiting core innovators with themselves at the time of survival. In other words, the control group is made by the same players and we can test whether exit can be explained by the fact that the given player has reduced its knowledge breadth or depth. We now consider core innovators that survived at least one period and exit at some point. We then compute the difference between their value of KB and KD and plot the distribution of these differences. The test becomes a one-sample t-test of paired differences with the null hypothesis that the data has zero mean and the alternative hypothesis that the data comes from a population with a mean greater than zero. Table 40 reports the results. The test confirms what found so far for diversification but rejects findings for knowledge accumulation. This means that exiting innovators had a significant lower value of diversification at the period before exit than in periods before survival.

**Table 40: Statistical tests for differences in the empirical cumulative distributions of knowledge breadth and depth between survival and exit groups**

| Two-sample Kolmogorov-Smirnoff Test |           |     |    |                       |         |              |                |        |        |        |        |
|-------------------------------------|-----------|-----|----|-----------------------|---------|--------------|----------------|--------|--------|--------|--------|
| Model                               | Variable  | Ns  | Ne | Result                | p-value | significance | test statistic | x*     | s*     | Fs(x*) | Fe(x*) |
| All core innovators                 | <b>KB</b> | 116 | 73 | H1: $F_S(x) < F_E(x)$ | 0.0002  | ***          | 0.299          | 0.6208 | 0.6208 | 0.3534 | 0.0548 |
| (all periods)                       | <b>KD</b> | 116 | 73 | H1: $F_S(x) < F_E(x)$ | 0.0000  | ***          | 0.378          | 0.000  | 0.0076 | 0.7069 | 0.3288 |

| Anderson-Darling Test (adjusted for ties) |           |     |    |                          |         |              |                 |                |  |  |  |
|---|-----------|-----|----|--------------------------|---------|--------------|-----------------|----------------|--|--|--|
| Model                                     | Variable  | Ns  | Ne | Result                   | p-value | significance | rank statistics | std. rank stat |  |  |  |
| All core innovators                       | <b>KB</b> | 116 | 73 | H1: $F_S(x) \neq F_E(x)$ | 0.0000  | ***          | 11.039          | 13.316         |  |  |  |
| (all periods)                             | <b>KD</b> | 116 | 73 | H1: $F_S(x) \neq F_E(x)$ | 0.0000  | ***          | 14.066          | 17.332         |  |  |  |

| One sample T-test of Paired Differences |           |    |        |        |          |   |         |              |        |    |  |
|---|-----------|----|--------|--------|----------|---|---------|--------------|--------|----|--|
| Survival-Exit                           |           |    |        |        |          | One-sample t-test of paired differences |         |              |        |    |  |
| Model                                   | Variable  | N  | Mean   | St.Dev | Skewness | st.error of difference                  | p-value | significance | t-stat | df |  |
| All core innovators                     | <b>KB</b> | 56 | 0.1006 | 0.326  | 0.204    | 0.3263                                  | 0.0124  | **           | 2.3068 | 55 |  |
| (all periods)                           | <b>KD</b> | 56 | 0.0001 | 0.010  | -3.570   | 0.0103                                  | 0.4659  |              | 0.086  | 55 |  |

#### A.5.4 Regression results using probit models

In the following, we report the analogous probit estimation of the logit models presented in the paper. The same findings discussed in the paper holds.

**Table 41: Pooled and panel data probit model estimation**

| PROBIT MODEL   | (1)                  | (2)                 | (3)                 | (4)                 |
|--|----------------------|---------------------|---------------------|---------------------|
|  | Pooled clust.VCE     | Pooled clust.VCE    | Panel Ran.Eff.      | Panel Ran.Eff.      |
|  | PC<24                | All obs.            | PC<24               | All obs.            |
| VARIABLES  | survival             | survival            | survival            | survival            |
| <b>Specialized</b>   |                      | -1.036<br>(0.968)   |                     | -2.180<br>(1.849)   |
| <b>KB bin category</b>                                       |                      |                     |                     |                     |
| KB bin 7 ( $0.6 \leq KB < 0.7$ )                             | 0.978***<br>(0.369)  |                     | 1.348**<br>(0.670)  |                     |
| KB bin 8 ( $0.7 \leq KB < 0.8$ )                             | 1.469**<br>(0.602)   |                     | 2.359**<br>(1.062)  |                     |
| KB bin 9 ( $0.8 \leq KB < 0.9$ )                             | 0.994**<br>(0.414)   |                     |                     |                     |
| <b>Restricted cubic spline (knots at KB = 0, 0.48, 0.89)</b> |                      |                     |                     |                     |
| r.c.s. KB1   |                      | -3.423<br>(2.940)   |                     | -7.643<br>(5.667)   |
| r.c.s. KB2   |                      | 5.213*<br>(2.892)   |                     | 10.44*<br>(5.903)   |
| <b>ln(KD)</b>  | -0.215<br>(0.248)    | -0.189<br>(0.251)   | -0.484<br>(0.389)   | -0.464<br>(0.383)   |
| <b>ln(KD)<sup>2</sup></b>                                    | -0.0159<br>(0.0142)  | -0.0147<br>(0.0144) | -0.0344<br>(0.0228) | -0.0336<br>(0.0225) |
| <b>PC</b>  | -0.00533<br>(0.0363) | 0.00272<br>(0.0142) | -0.0253<br>(0.0590) | 0.00156<br>(0.0347) |

| PROBIT   | (1)                 | (2)                 | (3)                 | (4)                 |
|--|---------------------|---------------------|---------------------|---------------------|
| MODEL  | Pooled clust.VCE    | Pooled clust.VCE    | Panel Ran.Eff.      | Panel Ran.Eff.      |
|  | PC<24               | All obs.            | PC<24               | All obs.            |
| VARIABLES  | survival            | survival            | survival            | survival            |
| <b>Experience</b> ( <i>base = new entrants</i> ) |                     |                     |                     |                     |
| 1 period   | 0.754***<br>(0.224) | 0.704***<br>(0.221) | -1.276**<br>(0.528) | -1.257**<br>(0.529) |
| 2 periods  | -0.612**<br>(0.277) | -0.548**<br>(0.269) | -1.389*<br>(0.716)  | -1.415*<br>(0.734)  |
| 3 periods  | -0.207<br>(0.357)   | -0.197<br>(0.350)   | -1.043<br>(0.824)   | -1.093<br>(0.838)   |
| 4 periods  | -0.363<br>(0.401)   | -0.283<br>(0.390)   | -1.444<br>(1.004)   | -1.352<br>(0.973)   |
| <b>Business cat.</b> ( <i>base = IDMs</i> )      |                     |                     |                     |                     |
| Fabless  | -0.154<br>(0.350)   | -0.0595<br>(0.361)  | -0.528<br>(0.820)   | -0.400<br>(0.818)   |
| Supplier   | 0.547<br>(0.455)    | 0.564<br>(0.438)    | 0.738<br>(0.787)    | 0.796<br>(0.798)    |
| Res.provider                                     | -0.433<br>(0.287)   | -0.429<br>(0.280)   | -0.799<br>(0.533)   | -0.762<br>(0.531)   |
| User   | -0.243<br>(0.244)   | -0.319<br>(0.240)   | -0.494<br>(0.492)   | -0.566<br>(0.496)   |
| <b>RK Persistence</b><br>(main trajet.)          | 9.502**<br>(4.251)  | 9.923**<br>(4.356)  | 15.43*<br>(8.535)   | 16.59*<br>(8.573)   |
| <b>Constant</b>                                  | -0.664<br>(1.243)   | 0.406<br>(1.441)    | -1.345<br>(1.904)   | 0.793<br>(2.492)    |
| Observations                                     | 286                 | 323                 | 286                 | 323                 |
| Log-Likelihood                                   | -132.6              | -136.3              | -129.7              | -133                |
| # of companies                                   |                     |                     | 122                 | 126                 |
| Sigma_u  |                     |                     | 1.276               | 1.336               |
| Rho  |                     |                     | 0.619               | 0.641               |
| Pseudo R2  | 0.189               | 0.216               |                     |                     |

**Notes:** Robust standard errors in parentheses (errors clustered by company for the pooled data)

Significance legend: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Vertical bars indicate joint significance of the variables reported to the left of the bar(s) (||| p<0.01, || p<0.05, | p<0.1). Business category 'Foundry' (in all models) and KB bins 2 and 10 (in Model 9) were omitted from the regressions as they predicted survival perfectly. For Model 9 only KB bins whose coefficient is statistically significant are reported.

**Table 42: Pooled and panel data probit model estimation with split sample**

| PROBIT    | (1)       | (2)       | (3)          | (4)      | (5)      | (6)          |
|-----------|-----------|-----------|--------------|----------|----------|--------------|
| MODEL     | Pooled    | Pooled    | Pooled       | Panel    | Panel    | Panel        |
|           | clust.VCE | clust.VCE | clust.VCE    | RandEff  | RandEff  | RandEff      |
|           | KB=0      | KB=0      | PC<24 & KB>0 | KB=0     | KB=0     | PC<24 & KB>0 |
| VARIABLES | survival  | survival  | survival     | survival | survival | survival     |

KB bin category (*base = bin 5*)

|                          |                     |                     |
|--------------------------|---------------------|---------------------|
| KB bin 6<br>(0.5≤KB<0.6) | 0.391<br>(0.332)    | 0.391<br>(0.332)    |
| KB bin 7<br>(0.6≤KB<0.7) | 1.075**<br>(0.472)  | 1.075**<br>(0.472)  |
| KB bin 8<br>(0.7≤KB<0.8) | 1.779***<br>(0.600) | 1.779***<br>(0.600) |

| PROBIT<br>MODEL                         | (1)<br>Pooled<br>clust.VCE<br>KB=0 | (2)<br>Pooled<br>clust.VCE<br>KB=0 | (3)<br>Pooled<br>clust.VCE<br>PC<24 & KB>0 | (4)<br>Panel<br>RandEff<br>KB=0 | (5)<br>Panel<br>RandEff<br>KB=0 | (6)<br>Panel<br>RandEff<br>PC<24 & KB>0 |
|---|------------------------------------|------------------------------------|--|---------------------------------|---------------------------------|---|
| VARIABLES                               | survival                           | survival                           | survival                                   | survival                        | survival                        | survival                                |
| KB bin 9<br>(0.8≤KB<0.9)                |                                    |                                    | 1.406**<br>(0.546)                         |                                 |                                 | 1.406**<br>(0.546)                      |
| ln(KD)                                  | 0.0194<br>(0.0673)                 | -0.470<br>(0.305)                  | 0.885<br>(0.657)                           | 0.0194<br>(0.0673)              | -0.470<br>(0.379)               | 1.094<br>(1.032)                        |
| ln(KD) <sup>2</sup>                     |                                    | -0.0267<br>(0.0169)                | 0.0451<br>(0.0390)                         |                                 | -0.0267<br>(0.0203)             | 0.0539<br>(0.0588)                      |
| PC                                      | 0.0302<br>(0.163)                  | 0.0561<br>(0.158)                  | -0.0576<br>(0.0361)                        | 0.0302<br>(0.161)               | 0.0561<br>(0.164)               | -0.0822<br>(0.0878)                     |
| <b>Experience</b> (base = new entrants) |                                    |                                    |  |                                 |                                 |   |
| 1 period                                | -0.963**<br>(0.377)                | -0.964**<br>(0.379)                | -0.570<br>(0.394)                          | -0.964***<br>(0.373)            | -0.964***<br>(0.371)            | -0.826<br>(0.907)                       |
| 2 periods                               | -1.073**<br>(0.421)                | -1.050**<br>(0.431)                | -0.164<br>(0.433)                          | -1.073***<br>(0.404)            | -1.050***<br>(0.405)            | -0.301<br>(0.751)                       |
| 3 periods                               | -1.101*<br>(0.566)                 | -0.932<br>(0.569)                  | 0.903<br>(0.647)                           | -1.102**<br>(0.530)             | -0.932*<br>(0.552)              | 0.968<br>(0.944)                        |
| 4 periods                               | -0.894<br>(0.600)                  | -0.926<br>(0.599)                  | 0.0352<br>(0.607)                          | -0.895<br>(0.597)               | -0.926<br>(0.602)               | -0.217<br>(1.157)                       |
| <b>Business cat.</b> (base = IDMs)      |                                    |                                    |  |                                 |                                 |   |
| Fabless                                 | -0.375<br>(0.456)                  | -0.359<br>(0.454)                  | -  | -0.375<br>(0.500)               | -0.360<br>(0.500)               | -                                       |
| Supplier                                | 0.359<br>(0.704)                   | 0.344<br>(0.714)                   | 0.454<br>(0.661)                           | 0.359<br>(0.654)                | 0.344<br>(0.654)                | 0.621<br>(1.114)                        |
| Res.provider                            | -0.00645<br>(0.353)                | 0.00900<br>(0.360)                 | -0.773**<br>(0.356)                        | -0.00647<br>(0.369)             | 0.00899<br>(0.370)              | -1.051<br>(0.955)                       |
| User                                    | -0.353<br>(0.324)                  | -0.318<br>(0.316)                  | -0.458<br>(0.349)                          | -0.353<br>(0.372)               | -0.318<br>(0.370)               | -0.669<br>(0.766)                       |
| <b>RK Persistence</b><br>(main trajet.) | 23.21***<br>(6.959)                | 22.45***<br>(6.764)                | 0.168<br>(6.562)                           | 23.21***<br>(8.031)             | 22.45***<br>(7.941)             | 1.880<br>(11.09)                        |
| <b>Constant</b>                         | -1.047<br>(1.138)                  | -3.149*<br>(1.644)                 | 4.964*<br>(2.941)                          | -1.047<br>(1.267)               | -3.149<br>(2.069)               | 6.238<br>(5.193)                        |
| Observations                            | 128                                | 128                                | 155  | 128                             | 128                             | 155                                     |
| Log-Likelihood                          | -66.73                             | -65.81                             | -54.53                                     | -66.73                          | -65.81                          | 83                                      |
| Sigma_u                                 |                                    |                                    |  | 0.00120                         | 0.00272                         | 0.899                                   |
| Rho                                     |                                    |                                    |  | 1.43e-06                        | 7.42e-06                        | 0.447                                   |
| Pseudo R2                               | 0.207                              | 0.218                              | 0.239                                      |                                 |                                 |   |

**Notes:** Robust standard errors in parentheses (errors clustered by company for the pooled data)

Significance legend: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Vertical bars indicates joint significance (||| p<0.01, || p<0.05, | p<0.1).

Business category 'Foundry' (in all models), 'Fabless' and KB bins 2 and 10 (in Model 3 and 6) are omitted from the regressions as they predict survival perfectly. For Models 3 and 6 only KB bins others than 2, 3 and 4 are reported. Bin is 5 used as baseline.



### A.5.5 Regression results using complementary log-log models

In the following, we report the analogous complementary log-log estimation of the logit models presented in the paper.

**Table 43: Pooled and panel data complementary log-log model estimation**

| C LOG-LOG<br>MODEL   | (1)<br>Pooled clust.VCE<br>PC<24 | (2)<br>Pooled clust.VCE<br>All obs. | (3)<br>Panel Ran.Eff.<br>PC<24 | (4)<br>Panel Ran.Eff.<br>All obs. |
|--|----------------------------------|-------------------------------------|--------------------------------|-----------------------------------|
| VARIABLES  | survival                         | survival                            | survival                       | survival                          |
| <b>Specialized</b>   |                                  | 2.206<br>(1.621)                    |                                | -2.180<br>(1.849)                 |
| <b>KB bin category</b>                                       |                                  |                                     |                                |                                   |
| KB bin 7 ( $0.6 \leq KB < 0.7$ )                             | -1.476**<br>(0.641)              |                                     | -1.76**<br>(0.869)             |                                   |
| KB bin 8 ( $0.7 \leq KB < 0.8$ )                             | -2.641**<br>(1.225)              |                                     | -3.35**<br>(1.421)             |                                   |
| KB bin 9 ( $0.8 \leq KB < 0.9$ )                             | -1.461**<br>(0.646)              |                                     | -2.33<br>(1.464)               |                                   |
| <b>Restricted cubic spline (knots at KB = 0, 0.48, 0.89)</b> |                                  |                                     |                                |                                   |
| r.c.s. KB1   |                                  | 7.470<br>(5.027)                    |                                | -7.643<br>(5.667)                 |
| r.c.s. KB2   |                                  | -10.53**<br>(5.182)                 |                                | 10.44*<br>(5.903)                 |
| <b>ln(KD)</b>  | 0.318<br>(0.372)                 | 0.293<br>(0.373)                    | 0.624<br>(0.500)               | -0.464<br>(0.383)                 |
| <b>ln(KD)<sup>2</sup></b>                                    | 0.0204<br>(0.0200)               | 0.0191<br>(0.0202)                  | 0.0428<br>(0.0290)             | -0.0336<br>(0.0225)               |
| <b>PC</b>  | -0.00760<br>(0.0615)             | -0.0189<br>(0.0350)                 | 0.0358<br>(0.0773)             | 0.00156<br>(0.0347)               |
| <b>Experience (base = new entrants)</b>                      |                                  |                                     |                                |                                   |
| 1 period   | 1.173***<br>(0.343)              | 1.107***<br>(0.339)                 | 1.589**<br>(0.638)             | -1.257**<br>(0.529)               |
| 2 periods  | 0.989**<br>(0.414)               | 0.952**<br>(0.404)                  | 1.700**<br>(0.862)             | -1.415*<br>(0.734)                |
| 3 periods  | 0.448<br>(0.549)                 | 0.492<br>(0.536)                    | 1.197<br>(1.019)               | -1.093<br>(0.838)                 |
| 4 periods  | 0.648<br>(0.572)                 | 0.613<br>(0.564)                    | 1.690<br>(1.237)               | -1.352<br>(0.973)                 |
| <b>Business cat. (base = IDMs)</b>                           |                                  |                                     |                                |                                   |
| Fabless  | 0.272<br>(0.429)                 | 0.182<br>(0.443)                    | 0.621<br>(1.007)               | -0.400<br>(0.818)                 |
| Supplier   | -0.851<br>(0.771)                | -0.816<br>(0.751)                   | -0.980<br>(1.033)              | 0.796<br>(0.798)                  |
| Res.provider   | 0.438                            | 0.487                               | 0.922                          | -0.762                            |

| C LOG-LOG<br>MODEL                      | (1)<br>Pooled clust.VCE<br>PC<24 | (2)<br>Pooled clust.VCE<br>All obs. | (3)<br>Panel Ran.Eff.<br>PC<24 | (4)<br>Panel Ran.Eff.<br>All obs. |
|---|----------------------------------|-------------------------------------|--------------------------------|-----------------------------------|
| VARIABLES                               | survival                         | survival                            | survival                       | survival                          |
|   | (0.375)                          | (0.363)                             | (0.662)                        | (0.531)                           |
| User                                    | 0.272<br>(0.340)                 | 0.387<br>(0.328)                    | 0.566<br>(0.603)               | -0.566<br>(0.496)                 |
| <b>RK Persistence</b><br>(main trajet.) | -16.60**<br>(6.688)              | -16.15**<br>(6.422)                 | -20.4**<br>(10.41)             | 16.59*<br>(8.573)                 |
| <b>Constant</b>                         | 1.175<br>(1.939)                 | -1.161<br>(2.262)                   | 1.537<br>(2.457)               | 0.793<br>(2.492)                  |
| Observations                            | 286                              | 323                                 | 286                            | 323                               |
| Log-Likelihood                          | -131.7                           | -135.2                              | -129.7                         | -133.1                            |
| # of companies                          |                                  |                                     | 122                            | 126                               |
| Sigma_u                                 |                                  |                                     | 1.466                          | 1.478                             |
| Rho                                     |                                  |                                     | 0.566                          | 0.571                             |
| Significance of rho                     |                                  |                                     | 0.023                          | 0.019                             |

**Notes:** Robust standard errors in parentheses (errors clustered by company for the pooled data)

Significance legend: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Vertical bars indicate joint significance of the variables reported to the left of the bar(s) (||| p<0.01, || p<0.05, | p<0.1). Business category 'Foundry' (in all models) and KB bins 2 and 10 (in Model 9) were omitted from the regressions as they predicted survival perfectly. For Model 9 only KB bins whose coefficient is statistically significant are reported.

**Table 44: Pooled and panel data complementary log-log model estimation with split sample**

| C LOG-LOG<br>MODEL                               | (1)<br>Pooled<br>clust.VCE<br>KB=0 | (2)<br>Pooled<br>clust.VCE<br>KB=0 | (3)<br>Pooled<br>clust.VCE<br>PC<24 & KB>0 | (4)<br>Panel<br>RandEff<br>KB=0 | (5)<br>Panel<br>RandEff<br>KB=0 | (6)<br>Panel<br>RandEff<br>PC<24 & KB>0 |
|--|------------------------------------|------------------------------------|--|---------------------------------|---------------------------------|---|
| VARIABLES  | survival                           | survival                           | survival                                   | survival                        | survival                        | survival                                |
| KB bin category ( <i>base = bin 5</i> )          |                                    |                                    |  |                                 |                                 |   |
| KB bin 6<br>(0.5≤KB<0.6)                         |                                    |                                    | -1.600**<br>(0.785)                        |                                 |                                 | -1.961<br>(1.347)                       |
| KB bin 7<br>(0.6≤KB<0.7)                         |                                    |                                    | -2.884**<br>(1.193)                        |                                 |                                 | -3.824<br>(2.669)                       |
| KB bin 8<br>(0.7≤KB<0.8)                         |                                    |                                    | -1.982**<br>(0.839)                        |                                 |                                 | -2.851<br>(2.426)                       |
| KB bin 9<br>(0.8≤KB<0.9)                         |                                    |                                    |  |                                 |                                 | -1.961<br>(1.347)                       |
| <b>ln(KD)</b>                                    | -0.00843<br>(0.0901)               | 0.603*<br>(0.331)                  | -1.558<br>(1.193)                          | -0.00842<br>(0.0826)            | 0.603<br>(0.408)                | -1.736<br>(1.487)                       |
| <b>ln(KD)<sup>2</sup></b>                        |                                    | 0.0330*<br>(0.0182)                | -0.0832<br>(0.0690)                        |                                 | 0.0330<br>(0.0216)              | -0.0854<br>(0.0853)                     |
| <b>PC</b>  | -0.0343<br>(0.191)                 | -0.0550<br>(0.176)                 | 0.0715<br>(0.0651)                         | -0.0343<br>(0.189)              | -0.0550<br>(0.186)              | 0.131<br>(0.142)                        |
| <b>Experience</b> ( <i>base = new entrants</i> ) |                                    |                                    |  |                                 |                                 |   |
| 1 period   | 1.492***                           | 1.514***                           | 0.913                                      | 1.492***                        | 1.514***                        | 1.338                                   |

| C LOG-LOG<br>MODEL                   | (1)<br>Pooled<br>clust.VCE<br>KB=0 | (2)<br>Pooled<br>clust.VCE<br>KB=0 | (3)<br>Pooled<br>clust.VCE<br>PC<24 & KB>0 | (4)<br>Panel<br>RandEff<br>KB=0 | (5)<br>Panel<br>RandEff<br>KB=0 | (6)<br>Panel<br>RandEff<br>PC<24 & KB>0 |
|--------------------------------------|------------------------------------|------------------------------------|--|---------------------------------|---------------------------------|---|
| VARIABLES                            | survival                           | survival                           | survival                                   | survival                        | survival                        | survival                                |
|                                      | (0.578)                            | (0.577)                            | (0.686)                                    | (0.574)                         | (0.575)                         | (1.489)                                 |
| 2 periods                            | 1.606**                            | 1.576**                            | 0.282                                      | 1.606***                        | 1.576**                         | 0.492                                   |
|                                      | (0.633)                            | (0.643)                            | (0.720)                                    | (0.608)                         | (0.615)                         | (1.264)                                 |
| 3 periods                            | 1.548**                            | 1.436**                            | -1.148                                     | 1.548**                         | 1.436**                         | -1.360                                  |
|                                      | (0.726)                            | (0.726)                            | (1.296)                                    | (0.697)                         | (0.723)                         | (1.500)                                 |
| 4 periods                            | 1.326*                             | 1.421*                             | 0.305                                      | 1.326*                          | 1.421*                          | 0.529                                   |
|                                      | (0.750)                            | (0.753)                            | (1.111)                                    | (0.765)                         | (0.776)                         | (1.761)                                 |
| Business cat. ( <i>base = IDMs</i> ) |                                    |                                    |  |                                 |                                 |   |
| Fabless                              | 0.479                              | 0.534                              | -0.516                                     | 0.479                           | 0.479                           | -                                       |
|                                      | (0.505)                            | (0.522)                            | (1.141)                                    | (0.612)                         | (0.612)                         |   |
| Supplier                             | -0.492                             | -0.437                             | 1.085*                                     | -0.492                          | -0.492                          | -0.910                                  |
|                                      | (1.205)                            | (1.211)                            | (0.596)                                    | (1.157)                         | (1.157)                         | (1.925)                                 |
| Res.provider                         | 0.0294                             | 0.0197                             | 0.621                                      | 0.0293                          | 0.0293                          | 1.617                                   |
|                                      | (0.430)                            | (0.439)                            | (0.596)                                    | (0.438)                         | (0.438)                         | (1.666)                                 |
| User                                 | 0.451                              | 0.389                              | 1.661                                      | 0.451                           | 0.451                           | 1.106                                   |
|                                      | (0.379)                            | (0.386)                            | (11.93)                                    | (0.415)                         | (0.415)                         | (1.284)                                 |
| <b>RK Persistence</b>                | -35.41***                          | -33.31***                          | -8.813                                     | -35.41***                       | -35.41***                       | -4.539                                  |
| (main traject.)                      | (11.53)                            | (10.75)                            | (5.522)                                    | (12.40)                         | (12.40)                         | (19.90)                                 |
| <b>Constant</b>                      | 1.396                              | 3.905**                            | 155  | 1.396                           | 1.396                           | -10.29                                  |
|                                      | (1.559)                            | (1.919)                            | -54.91                                     | (1.735)                         | (1.735)                         | (7.596)                                 |
| Observations                         | 128                                | 128                                | -0.516                                     | 128                             | 128                             | 155                                     |
| Log-Likelihood                       | -65.79                             | -64.65                             | (1.141)                                    | -65.79                          | -64.65                          | -54.68                                  |
| # of companies                       |                                    |                                    |  | 84                              | 84                              | 83                                      |
| Sigma_u                              |                                    |                                    |  | 0.000740                        | 0.00173                         | 1.573                                   |
| Rho                                  |                                    |                                    |  | 3.33e-07                        | 1.83e-06                        | 0.601                                   |
| Significance or rho                  |                                    |                                    |  | 1                               | 0.498                           | 0.251                                   |

**Notes:** Robust standard errors in parentheses (errors clustered by company for the pooled data)

Significance legend: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Vertical bars indicates joint significance (||| p<0.01, || p<0.05, | p<0.1).

Business category 'Foundry' (in all models), 'Fabless' and KB bins 2 and 10 (in Model 3 and 6) are omitted from the regressions as they predict survival perfectly. For Models 3 and 6 only KB bins others than 2, 3 and 4 are reported. Bin is 5 used as baseline.

## A.5.6 Scatterplot of core innovators in the KB-KD space

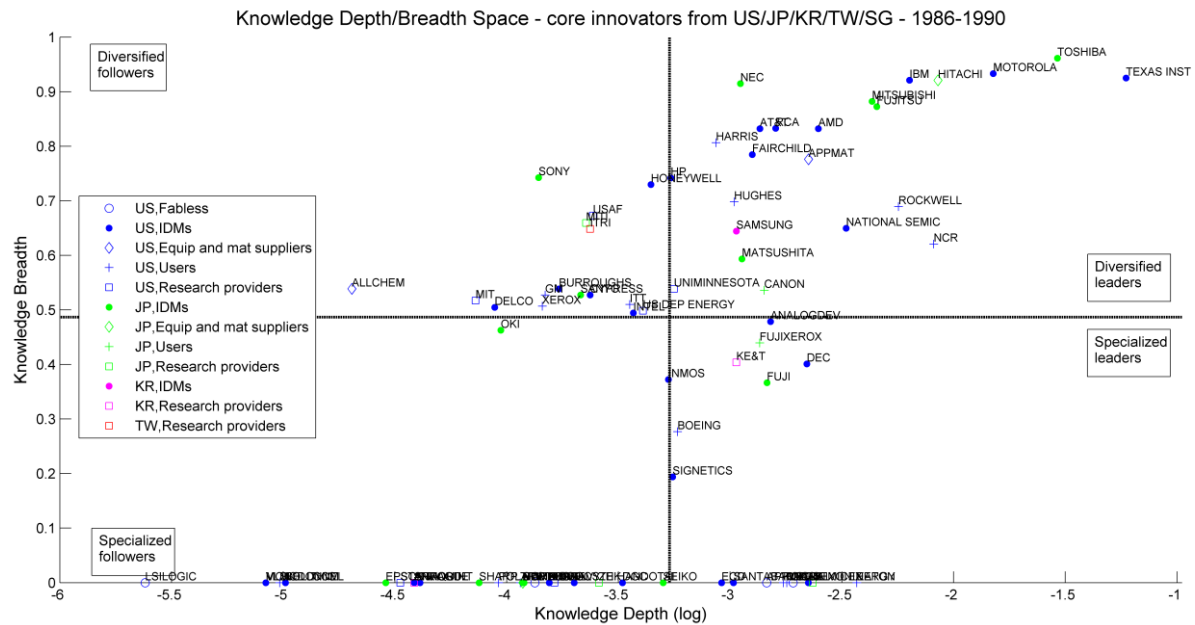


Figure 57: Core innovators' location in the knowledge breadth and depth space between 1986 and 1990

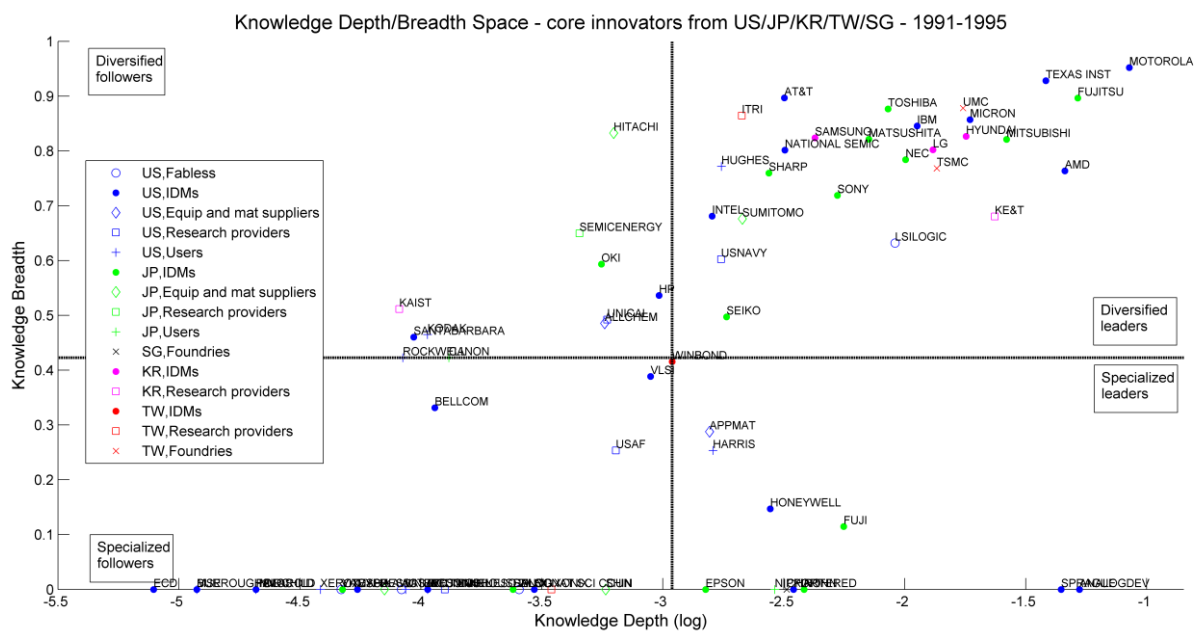


Figure 58: Core innovators' location in the knowledge breadth and depth space between 1991 and 1995

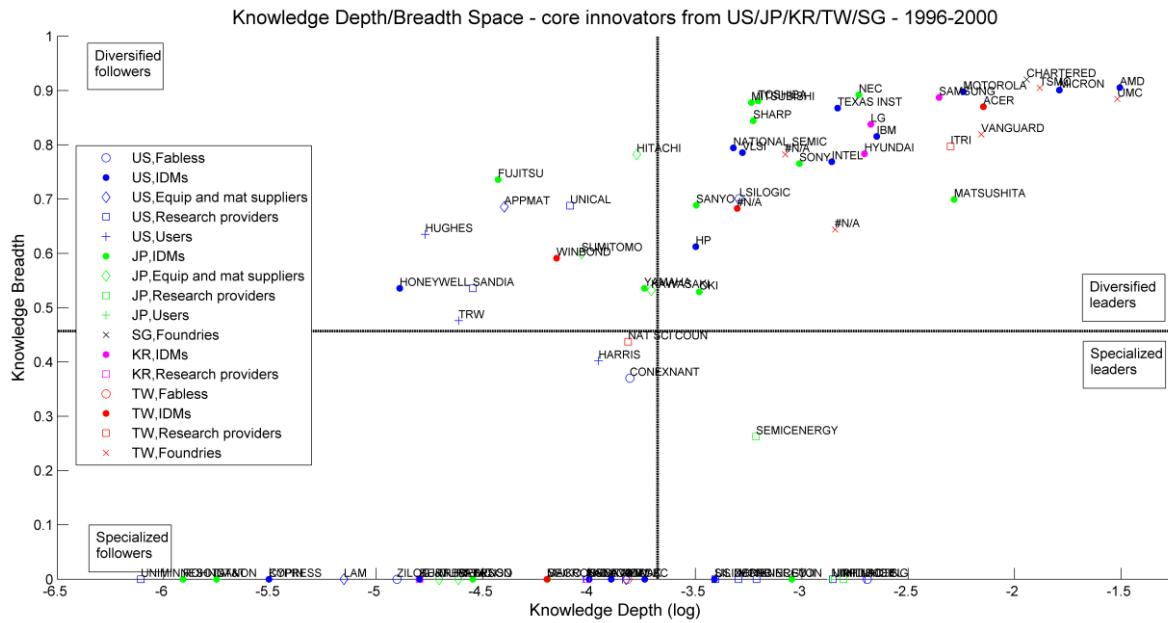


Figure 59: Core innovators' location in the knowledge breadth and depth space between 1996 and 2000

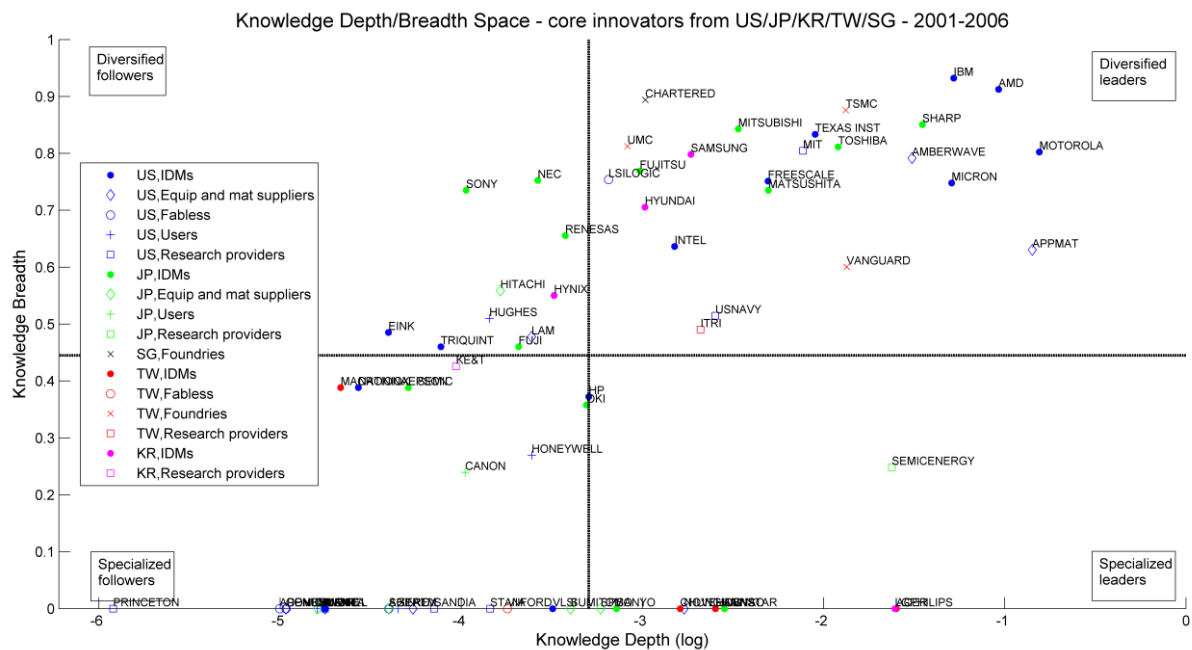


Figure 60: Core innovators' location in the knowledge breadth and depth space between 2001 and 2006

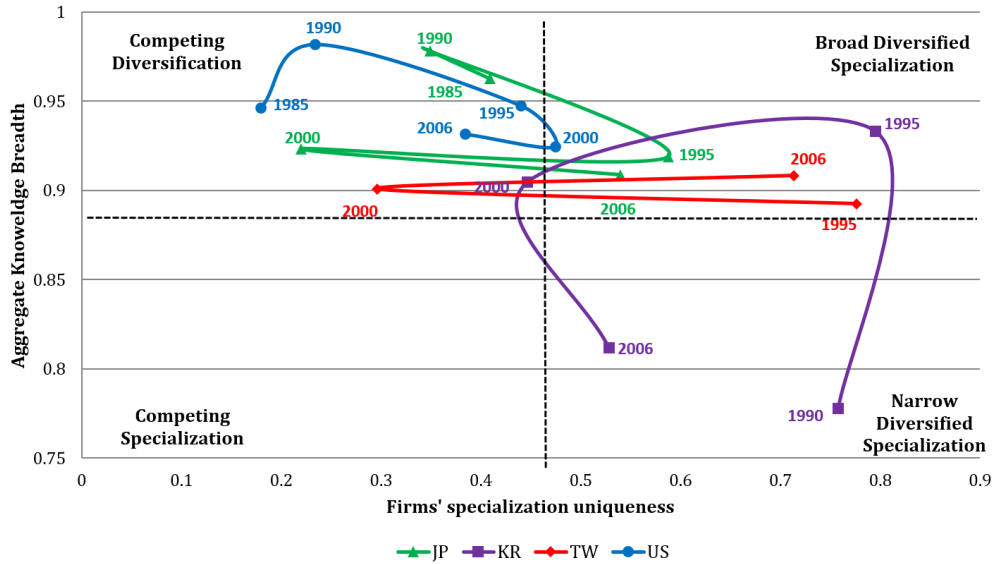
### A.5.7 The linkage between micro- and macro-diversification strategies

In this section, we present the results of the analysis of how countries achieve technical knowledge diversification. Aggregate diversification can be the results of two different scenarios. At any given level, aggregate knowledge breadth, can be achieved through the sum of firms' individual non-overlapping specialization patterns in different domains or through pooling together partly or fully overlapping firms' areas of specialization. Figure 61 shows the sources of aggregate technical knowledge diversification for S.Korea, Taiwan, Japan and US. On the vertical axis we report values of aggregate knowledge breadth, computed using the same formula intro-

duced in equation 5.9, where the knowledge genetic heritage of a country in a given domain is calculated using all patents granted to firms from the given country. The horizontal axis reports values of the knowledge uniqueness index. This index measures to which extent aggregate knowledge breadth is the result of overlapping firms' specialization patterns. The index takes a value of one (i.e. full knowledge uniqueness at the firm level) when a country's knowledge in each of the domains in which it is active is entirely concentrated within one firm. It takes a value of zero when in each domains knowledge is equally spread across firms. The index is calculated as follows.

$$Uniqueness - Within_c = \frac{1}{K} \cdot \sum_{j=1}^K \left( 1 - \frac{1 - \sum_{i=1}^{N_c} \left( \frac{proximity_{ij}}{\sum_{i=1}^{N_c} proximity_{ij}} \right)^2}{1 - \left( \frac{1}{N_c} \right)^2 \cdot N_c} \right) \quad (A.5.1)$$

Where,  $C$  is the country index,  $i$  goes from 1 to the number of firms in country  $C$  ( $N_c$ ),  $J$  goes from 1 to the number of technology domains ( $K$ ) and  $proximity_{ij}$  is the amount of genetic proximity of firm  $i$  in area  $j$ , computed using equation 5.8. It is worth noting how the uniqueness index is normalized by the maximum possible value attainable if the sum of a country's genetic proximity in each given domain would be randomly distributed across firms from the given country. Therefore it effectively control for the number of patenting entities in a country.



**Figure 61: Countries' knowledge diversification paths**

Dashed lines in Figure 61 are drawn to mark the average values of aggregate knowledge breadth and uniqueness across countries and time. The analysis highlights a big difference in the way former leaders (US and Japan) and successful latecomers (S.Korea and Taiwan) achieved technical knowledge diversification. US and Japan mainly followed a competing specialization pattern, with large aggregate knowledge breadth attained through overlapping firms knowledge, even though Japan was diversily specialized in in the first half of the 1990s and of the 2000s. In contrast, Taiwan evolved from a narrow diversification pattern to a broad diversified specializa-

tion, passing through a competing diversification stage. However, they are always located around the average boundaries of the space, therefore, we cannot exclude the possibility that this specialization pattern may be random. Korea has a clearer trajectory instead. They start from narrow diversified specialization, then evolved to broad diversified specialization but then reduced their knowledge breadth and uniqueness a little. Singapore is not reported due to the limited number of firms within the country.





# 9 VALORISATION

In accordance with article 23.5 of the Regulation governing the attainment of doctoral degrees at Maastricht University, this addendum discusses the valorization opportunities presented by this PhD thesis. I will first discuss the degree of innovativeness of the theories and methods presented here and then elaborate on the socio-economic relevance of this work for several target groups.

## 9.1 Degree of innovativeness

This thesis presented a number of novel insights about technology dynamics and latecomers' technological catching-up by developing a new theoretical and methodological framework to study the evolution of technology and firms' inventive strategies. The theoretical contribution of this thesis is the introduction of a new way of looking at catching up and technology dynamics that is grounded on the interaction between prevailing and emerging engineering problems and the variety of approaches to solve them. We also argued how changes in problem-solving approaches can lead to solve technical bottlenecks and spark the rise of new products. Ultimately, this can lead to the emergence of new engineering challenges that need to be tackled. In the thesis, we also developed a theoretical framework that links the level of technical knowledge modularity in an industry, the extent to which technical change is knowledge replacing and the urge of knowledge diversification for survival at the technological frontier.

This dissertation also contributes new methods to identify changes in engineering design trajectories, i.e. in the way engineering problems are solved and how their importance is perceived relative to other problems. We also introduced an algorithm to classify inventions, firms (and, potentially, regions and countries) with respect to the exploitation-exploration spectrum of problem-solving approaches and their focus on prevailing or emerging engineering problems. In addition, we designed a methodological framework to identify technology domains and assess their life cycle stage. Another main methodological contribution of this thesis lays in the creation of an index of specialization at the country level that provides a more realistic micro-founded picture of inventing activities by firms in a country than previously available indices. Finally, we also developed a model to predict a firm's probability of persisting innovating at the frontier given its level of knowledge breadth, depth and the extent to which past knowledge is useful today.

## 9.2 Socio-economic relevance

The theoretical and methodological framework presented in this thesis has important implications for several target groups. First and foremost, they are the first step of a rich research agenda whose ambitious goal is to quantitatively study technology evolution at the level of engineering problems and problem-solving approaches. We argued in the thesis how this may be the key to understand and, possibly, predict technology dynamics that lead to clustering of big innovations in time and contributes to the formation of economic cycles. This area of research is potentially very fertile and has been explored only partially. The rich admixture of concepts, theories and methods from evolutionary economics, complexity studies and strategy, which is found in this dissertation, also testifies how this new direction of theoretical thinking could affect several fields. For instance, one could use the same theoretical and methodological framework to analyse the knowledge evolution of specific scientific domains of interest, using publication instead of patent data. A unifying theory of knowledge evolution could be searched. Such theory would be based on the central notions defined in this thesis, namely that science and technology advance by solving problems and that problems are tackled by a variety of approaches. It would then need to be integrated with a sociological explanation of why the variety of approaches changes over time. Such theory could then be validated with data by applying the methods developed in this thesis.

Secondly, the theoretical and methodological contributions of this dissertation should be of interest to technology development practitioners. In fact, the research agenda that this thesis started can potentially lead to commercially viable and empirically grounded consultancy services in the field of strategy and business intelligence for technology development. A number of similar consultancy companies have been developed in the last few years attempting to consult firms about their location in the technology landscape. To the best of the author's knowledge (which is fairly limited in this area), none of these services are based on a dynamic perspective on technology. Rather, they take the technology space as given and consult firms on their current position. In contrast, the notions of a constantly evolving structure of the system of engineering problems and of the changing variety of existing problem solving approaches provide a much more accurate way of describing technology dynamics. This is enormously more appealing for business practitioners as it can provide detailed suggestions on the firm's position in the prevailing engineering trajectories and consult on the existence of alternative approaches. The core of such consultancy service could be based on the algorithms developed in this thesis to identify the main paths of engineering improvements in a given industry, detect technology domains, assess their life cycle stage, reveal companies' comparative advantage across technology domains and predict the probability that a company persist innovating at the technology frontier. Further refinement of these algorithms would focus on converting them in predictive tools to forecast the probability of an upcoming change in trajectory and assess how the firm is positioned in the technology space to take advantage or react to that. An additional service could address the problem of identifying useful external sources of knowledge. This could take the form of suggesting possible partners, acquisition targets, or individual inventors to hire, that have expertise in the new design trajectory or in emerging technology domains.

Similarly, the same kind of analysis could be of great interest to regional policy makers. The strength of a region's knowledge base in the industry of choice could be evaluated given the direction of technology evolution in terms of the current prevailing engineering problems and problem solving approaches, and the position of the region's firms or research lab in the technology space. High-resolution policy suggestions could be derived on how to steer the evolution

of the region's knowledge base toward the desired targets. Based on our analysis emerging areas of the technology space could be detected and targeted to build an early comparative advantage. Tailor-made subsidies to incentivize research and development effort and collaborations across firms in a given technology domain could be designed. For instance, if a region lacks any inventive activity in an emerging area of research in an industry that used to be strong in the region, one could identify possible external partners for collaborations based on knowledge complementarity. Alternatively, our method could also identify possible efficient research avenues, for firms that are currently in the region, to navigate the technology space until the desired emerging area is reached.

The pursuit of the commercialization avenues sketched above is not the author's current priority. However, being him a scholar interested in evolution, he recognizes that preferences, opportunities and paths can change in ways that are sometimes difficult to foresee. Yet, at the moment, the author intends to focus mostly on accomplishing the rich research agenda that started with this thesis and disseminating the results prevalently within academia. Perhaps this will actually lead to build a stronger recognition in this field that would facilitate future possible consultancies.

# 10 ABOUT THE AUTHOR

Giorgio Triulzi was born in 1984 in Milan, Italy. He is a postdoctoral fellow at the Institute of Data, System and Society at the Massachusetts Institute of Technology and an affiliated researcher at UNU-MERIT. His research interests focus on understanding drivers, direction and effects of technological change by applying theories and methods from complex system analysis, evolutionary economics and strategy.

Giorgio previously held positions as postdoctoral research assistant at the International Design Centre of Singapore University of Technology and Design and MIT, part-time lecturer at the Maastricht School of Business and Economics (SBE), PhD fellow at UNU-MERIT and at SBE. Giorgio has a PhD in *Economics and Policy Studies of Technical Change* from UNU-MERIT and Maastricht University, a Master of Science (*cum laude*) degree in *Economics of International Markets and New Technologies* from Bocconi University and a Bachelor of Science degree in the same field from the same institution. He has spent visiting stays as junior research assistant (pre-doctoral level) at the Institute of Institutional and Innovation Economics at the University of Bremen, as master student at the Eindhoven University of Technology (Netherlands) and as undergraduate student at the Universite' Louis Pasteur (France, now University of Strasbourg).



UNITED NATIONS  
UNIVERSITY

**UNU-MERIT**

## 2015

### 95. Giorgio Triulzi

Looking for the Right Path: Technology Dynamics, Inventive Strategies and Catching-up in the Semiconductor Industry

### 94. Abdul Baseer Qazi

Knowledge flows and networks in the ICT sector: The case of Pakistan

### 93. Ajay Thutupalli

Technology Paradigm Shifts in Agriculture: Drivers of Sustainability and Catch up

### 92. Eduardo Urias

Improving access to HIV/AIDS treatment in Brazil: When are compulsory licenses effective in price negotiations?

### 91. Francesca Guadagno

Why have so few countries industrialised?

### 90. Daniel Opolot

The evolution of beliefs and strategic behavior

### 89. Alejandro Lavopa

Structural Transformation and Economic Development: Can Development Traps be Avoided?

### 88. Jinjin Zhao

Urban water management reform: The case of China

## 2014

### 87. Dirk Crass

The Impact of Brands on Innovation and Firm Performance: Empirical Evidence from Germany

### 86. Samyukta Bhupatiraju

The Geographic Dimensions of Growth and Development

### 85. François Lafond

The evolution of knowledge systems

### 84. Annalisa Primi

Promoting Innovation in Latin America: What Countries Have Learned (and What They Have Not) in Designing and Implementing Innovation and Intellectual Property Policies

### 83. Fatoumata Lamarana Diallo

Evaluation of Meal and Deworming Programs for Primary Schools in Rural Senegal

## 2013

### 82. Anant Kamath

Information Sharing through Informal Interaction in Low-Tech Clusters

### 81. Flavia Pereira de Carvalho

What we talk about when we talk about Brazilian Multinationals: an investigation on Brazilian FDI, economic structure, innovation and the relationship between them

### 80. Jun Hou

Complementarity in Innovation and Development: A Cross-country Comparison

### 79. Rufin Baghana

Impacts of Government Incentives to R&D, Innovation and Productivity: A Microeconometric Analysis of the Québec Case

### 78. Lilia I. Stubrin

High-Tech Activities in Emerging Countries: A Network perspective on the Argentinean biotech activity

## 2012

### 77. Abdul Waheed

Innovation Determinants and Innovation as a Determinant: Evidence from Developing Countries

### 76. Bilal Mirza

Energy Poverty and Rural Energy Markets in Pakistan

### 75. Benjamin Engelstätter

Enterprise Software and Video Games: An Empirical Analysis

**Fulvia Farinelli**

Natural Resources, Innovation and Export Growth:  
The Wine Industry in Chili and Argentina

**Rodolfo Lauterbach**

Innovation in Manufacturing: From Product Variety  
and Labor Productivity Growth to Economic  
Development in Chile

**74. Kirsten Wiebe**

Quantitative Assessment of Sustainable  
Development and Growth in Sub-Saharan Africa.

**73. Julio Miguel Rosa**

Organizational Strategies, Firms' Performance and  
Spatial Spillovers. The Canadian Case in Research  
and Development.

**Johannes Wilhelmus Marie Boels**

Joseph Schumpeter, honderd jaar economische  
ontwikkeling. Een historisch-theoretische  
beschouwing.

## 2011

**72. Daniel Vertesy**

Interrupted Innovation: Emerging economies in the  
structure of the global aerospace industry.

**71. Tina Saebi**

Successfully managing alliance portfolios: an  
alliance capability view.

**70. Nora Engel**

Tuberculosis in India - A case of innovation and  
control.

**69. Evans Mupela**

Connectivity and growth in Sub-Saharan Africa: The  
role of communication satellites

**68. Nantawan Kwanjai**

Cross cultural intelligence amid intricate cultural  
webs: A tale of the UnDutchables in the land of  
1002 smiles

**67. Lina Sonne**

Innovation in Finance to Finance Innovation:  
Supporting pro-poor entrepreneur-based  
innovation

## 2010

**66. Fernando Santiago**

Human Resources Management Practices and  
Learning for Innovation in Developing Countries:  
Pharmaceutical Firms in Mexico

**65. Zakaria Babutsidze**

Essays on Economies with Heterogenous Interacting  
Consumers

**64. Bertha Vallejo**

Learning and Innovation Under Changing Market  
Conditions: The Auto Parts Industry in Mexico

**63. Donatus Ayitey**

Technical Change, Competitiveness and Poverty  
Reduction: A Study of the Ghanaian Apparel  
Industry

**62. Sergey Fillipov**

Multinational Subsidiary Evolution: Corporate  
Change in New EU Member States

**61. Asel Doranova**

Technology Transfer and Learning under the Kyoto  
regime; Exploring the Technological Impact of CDM  
projects in developing countries

## 2009

**60. Alexis Habiyaremye**

From Primary Commodity Dependence to  
Diversification and Growth". "Absorptive Capacity  
and Technological Catch Up in Botswana and  
Mauritius".

**59. Yoseph Getachew**

The Role of Public Capital in Economic Development

**58. Sandra Leitner**

Embodied Technological Change and Patterns of  
Investment in Austrian Manufacturing

**57. Semih Akçomak**

The Impact of Social Capital on Economic and Social  
Outcomes

**56. Abraham Garcia**

The Role of Demand in Technical Change

**55. Saurabh Arora**

Coherence in socio-technical systems: a network perspective on the innovation process

**2008**

**54. Rutger Daems**

Medicines for the developing world

**53. Johannes Hanel**

Assessing Induced Technology: Sombart's Understanding of Technical Change in the History of Economics

**52. Rifka Weehuizen**

Mental Capital: the economic significance of mental health

**51. Danielle Cloudt**

The relationship between R&D partnership formation, social embeddedness and innovative performance

**50. Sabine Fuss**

Sustainable Energy Development under Uncertainty

**2007**

**49. Tobias Kronenberg**

Reconciling Environmental Conservation with Economic Prosperity: The Feasibility of Double Dividends in the Short and Long Run

**48. Viktoria Kravtsova**

Assessing the Impact of Foreign Direct Investment in Transition Economies

**47. Suhail Sultan**

The Competitive Advantage of Small and Medium Sized Enterprises: The Case of Jordan's Natural Stone Industry

**2006**

**46. Bulat Sanditov**

Essays on Social Learning and Imitation

**45. Mamata Parhi**

Dynamics of New Technology Diffusion: A Study of the Indian Automotive Industry

**44. Andreas Reinstaller**

Social structures and the innovation process: Their role in the demand of firms and consumers

**43. Rose Kiggundu**

Innovation systems and development: the journey of a Beleaguered Nile Perch Fishery in Uganda

**42. Thomas Pogue**

The Evolution of Research Collaboration in South African Gold Mining: 1886-1933

**41. Geoffrey Gachino**

Foreign Direct Investment, Spillovers and Innovation: The Case of Kenyan Manufacturing Industry

**40. Önder Nomaler**

Technological Change, International Trade and Growth: An Evolutionary, Multi-Agents-Based Modeling Approach

**2005**

**39. Samia Satti Osman Mohamed-Nour**

Change and Skill Development in the Arab Gulf Countries

**38. Elad Harison**

Intellectual Property Rights: Economics and Policy Analysis

**37. Daniel Dalohoun**

The relationship between R&D partnership formation, social embeddedness and innovative performance: a multi-level approach of social embeddedness

**36. Müge Özman**

Networks, Organizations and Knowledge

**35. Bas Straathof**

Product variety and economic growth: The counteracting effects of scale and idiosyncrasy

**34. Wilfred Schoenmakers**

Knowledge Flows between Multinational Companies: A Patent Data Analysis

**33. Myriam Cloudt**

Mergers and Acquisitions (M&As) in High-Tech Industries: Measuring the Post-M&A Innovative Performance of Companies

## 2004

### 32. Paola Criscuolo

R&D Internationalisation and Knowledge Transfer. Impact on MNEs and their Home Countries

### 31. Maarten Verkerk

Trust and Power on the Shop Floor

### 30. Gottfried Leibbrandt

Adoption, harmonization and succession of network technologies across countries

### 29. Mark Sanders

Skill Biased Technical change - Its Origins, the Interaction with the Labour Market and Policy Implications

## 2003

### 28. Nadine Roijakkers

Inter-firm cooperation in high-tech industries: a study of R&D partnerships in pharmaceutical biotechnology

### 27. Viki Sonntag

Speed, Scale and Sustainability

### 26. Masaru Yarime

From End-of-Pipe Technology to Clean Technology

### 25. Stéphane Malo

The combinatorial Chemistry Revolution - Sustaining a Superior Performance Position through Technological Learning

## 2002

### 24. Annelies Hogenbirk

Determinants of Inward Foreign Direct Investment: the Case of the Netherlands

## 2001

### 23. John Adeoti

Technology Investment in Pollution Control in Sub-Saharan Africa: The Case of the Nigerian Manufacturing Industry

### 22. Edward Huizenga

Innovation Management: How Frontrunners Stay Ahead. An Empirical Study on Key Success Factors in the ICT sector

## 2000

### 21. Machiel van Dijk

Technological Change and the Dynamics of Industries. Theoretical Issues and Empirical evidence from Dutch Manufacturing

## 1999

### 20. Jan Cobbenhagen

Managing Innovation at the Company Level: A Study on Non-Sector-Specific Success Factors

### 19. Marjolein Caniëls

Regional Growth Differentials: The Impact of Locally Bounded Knowledge Spillovers

## 1998

### 18. Aldo Geuna

Resource allocation and knowledge production: Studies in the economics of university research

## 1996

### 17. Reinoud Joosten

Dynamics, Equilibria, and Values

### 16. Hugo Kruiniger

Investment, R&D, and the Financing Decisions of the Firm

## 1995

### 15. Hans van Meijl

Endogenous Technological Change: The Case of Information Technology. Theoretical Considerations and Empirical Results

### 14. René Kemp

Environmental Policy and Technical Change. A Comparison of the Technological Impact of Policy Instruments

### 13. Rohini Acharya

The Impact of New Technologies on Economic Growth and Trade. A Case Study of Biotechnology



**12. Geert Duysters**

The Evolution of Complex Industrial Systems. The Dynamics of Major IT Sectors

**11. Marjan Groen**

Technology, Work and Organisation, A Study of the Nursing Process in Intensive Care Units

**1994**

**10. Huub Meijers**

On the Diffusion of Technologies in a Vintage Framework; Theoretical Considerations and Empirical Results

**9. Theon van Dijk**

The Limits of Patent Protection. Essays on the Economics of Intellectual Property Rights

**8. Hans Voordijk**

Naar Integrale Logistiek in Bedrijfsketens, Ontwikkelingen in de Bouw

**1993**

**7. Paul Diederer**

Technological Progress in Enterprises and Diffusion of Innovations. Theoretical Reflections and Empirical Evidence.

**6. Ben Dankbaar**

Economic Crisis and Institutional Change. The crisis of Fordism from the perspective of the automobile industry

**5. Hanno Roberts**

Accountability and Responsibility: The Influence of Organisation Design on Management Accounting

**1992**

**4. Bart Verspagen**

Uneven Growth Between Interdependent Economies. An Evolutionary View on Technology Gaps, Trade and Growth

**3. Sjoerd Romme**

A Self-organization Perspective on Strategy Formation

**1989**

**2. John Spangenberg**

Economies of Scale, and Atmosphere in Research Organisations

**1988**

**1. John Hagedoorn**

Evolutionary and heterodox innovation analysis: a study of industrial and technological development in process control and information technology