Empirical and Theoretical Observations in Trade Secrecy: Statutory Prescriptions and Endogenous Growth

by

Mackey Craven

S.B. Biological Engineering
Massachusetts Institute of Technology, 2010

S.B. Mathematics
Massachusetts Institute of Technology, 2010

Submitted to the Engineering Systems Division in partial fulfillment of the requirements for the degree of

Master of Science in Technology and Policy at the
Massachusetts Institute of Technology
September 2010

© 2010 Massachusetts Institute of Technology. All rights reserved.

Signature of author:

Mackey Craven
Technology and Policy Program
August 18, 2010

Certified by:

Professor Fiona E. Murray
Sarofim Family Career Development Professor
Associate Professor of Management of Technological Innovation & Entrepreneurship
Thesis Supervisor

Accepted by:

Professor Dava J. Newman
Professor of Aeronautics and Astronautics and Engineering Systems
Director of Technology and Policy Program
MacVicar Fellow
Empirical and Theoretical Observations in Trade Secrecy: Statutory Prescriptions and Endogenous Growth

by

Mackey Craven

Submitted to the Engineering Systems Division on August 18, 2010 in partial fulfillment of the requirements for the degree of Master of Science in Technology and Policy

Abstract

Recent scholarship has suggested that trade secrecy may be as important, if not more important, than patenting as an intellectual property protection mechanism in the United States economy. While patent protection has always been an institution in the United States as a result of its inclusion in the Constitution, trade secret protection has only been statutorily recognized by the federal government in the last half-century. Similarly, while there has been extensive theoretical and empirical research on the incentives patent protection creates for firms to innovate, and the resulting effects on economic growth, there have been comparatively few theoretical studies and only a handful of empirical papers investigating the effects of trade secrecy. This leaves policy makers with little to no understanding of the effects of trade secrecy on the economy or how to influence trade secrecy through policy.

This thesis provides insight into the second question by investigating the effect of implementing statutory trade secret law on patent application and trade secrecy litigation rates. Specifically, this thesis performs a difference in difference analysis on the implementation of the Pennsylvania Uniform Trade Secrets Act with respect to the above metrics. While the effect of trade secret litigation is inconclusive, this thesis finds that patent application rates and the rate of growth of patent applications decrease in response to the implementation of statutory trade secret law. This implies a theorized, but never before measured, substitution effect between patents and trade secrets. In addition, this shows that the level of trade secret protection can be influenced by policymakers through statutory law. When combined with an empirical study on the welfare effects of trade secret protection, this thesis will give policymakers a reference point from which to consider the benefits of further statutory trade secret law.

Thesis Supervisor: Fiona E. Murray
Sarofim Family Career Development Professor
Associate Professor of Management of Technological Innovation & Entrepreneurship
Acknowledgements

It is with the utmost gratitude that I thank Prof. Fiona Murray for her involvement in my academic and professional life over the last eight months. Her mentorship and guidance have helped me in more ways than can be acknowledged here. Nonetheless, I would like to highlight her most significant contributions. Prof. Murray not only decided to hire me as a research assistant while I was still an undergraduate, but gave me a great deal of freedom with which to choose a general topic for my thesis and develop the study herein. Her trust in my decisions and encouragement to engage the greater research community has been invaluable. In addition, her flexibility with respect to my unorthodox work schedule was instrumental to my success throughout my time at the Technology and Policy Program.

Without the incredible support of Sydney Miller, I would not have been able to successfully navigate my final year at MIT and complete my degrees as intended. From helping me associate my courses with the correct degree applications to providing flexibility in the completion date of this thesis, Sydney has been an invaluable resource. Beyond that, her kindness and understanding over the past year significantly enhanced my experience in the Technology and Policy Program.

Without the willingness of Dr. Frank Field to keep an open mind as I made my case for admission to the Technology and Policy Program, I certainly would not have had the opportunity to write this thesis. His belief in my ability to succeed has been inspiring.

I thank Prof. Dava Newman and the admissions committee of the Technology and Policy Program for accepting my unorthodox application and allowing me to make the most of my fourth year at MIT. The number of amazing people I have met and opportunities I have had as a direct result of their decision has been astounding.

I acknowledge Jose Pacheco for his assistance in securing the funding for my research assistantship and in doing so, allowing me to graduate with my Bachelor’s degrees as planned in June 2010.

I sincerely thank the entire staff of the MIT Entrepreneurship Review, with special thanks to Bill Aulet and Erdin Beshimov, for connecting me with the MIT Entrepreneurship Center and my advisor Prof. Fiona Murray. Without the relationships I fostered while a founding editor in the organization, I would not have had the opportunity to write this thesis.

I acknowledge Prof. Fiona Murray’s research group, with special to Michael Bikard and Avi Wolfson, for their insightful comments about my proposed research design.

I thank Prof. Joshua Gans for providing valuable criticism of my original research design and for being a sounding board for ideas for future work.

I am eternally grateful to my parents, Dr. Laurie Lipman and Alex Craven, as well as my stepmother, Hannah Craven, for emphasizing and supporting my education throughout my life and giving me the opportunity to study at the university of my choice without regard to financial cost.

I give my deepest thanks to Vanessa Gusick for mentally and emotionally supporting me throughout my four years at MIT. Her love, kindness and understanding have been a source of stability and comfort in an otherwise hectic, and occasionally stressful, academic environment.
Table of Contents

List of Tables ................................................................................................................................... 9
List of Figures ..................................................................................................................................... 10

Chapter 1: Introduction to the Economics of Trade Secrecy ............................................................. 11
  1.1 Endogenous Growth and Codified Knowledge ........................................................................... 11
  1.2 Creation of Codified Knowledge ............................................................................................... 14
  1.3 (Non)-Disclosure of Codified Knowledge .................................................................................. 16
  1.4 Trade Secrecy Defined ............................................................................................................... 20
  1.5 Choosing Between Patent and Trade Secret Protection .......................................................... 21
  1.6 Endogenous Growth and Trade Secrecy ................................................................................... 24
  1.7 Importance of Trade Secrecy in the U.S. Economy ............................................................... 27

Chapter 2: Trade Secret Law ............................................................................................................... 29
  2.1 History of Trade Secret Law in the United States ..................................................................... 29
  2.2 Trade Secret Law as Statutory Intellectual Property Law ....................................................... 39

Chapter 3: Empirical Design ................................................................................................................ 42
  3.1 The Uniform Trade Secrets Act: A Natural Experiment .......................................................... 42
  3.2 Observing Trade Secret Litigation and Patent Applications ................................................... 44
  3.3 The Pennsylvania Uniform Trade Secrets Act ........................................................................... 47
  3.4 Adaptation of the Basic Difference in Difference Model ........................................................... 49
  3.5 Expected Results ....................................................................................................................... 53

Chapter 4: Empirical Methodology ..................................................................................................... 57
  4.1 Industry Sector Selection and Refinement .................................................................................. 57
  4.2 Data Acquisition and Difference in Difference Regressions ..................................................... 59

Chapter 5: Results .................................................................................................................................. 63
  5.1 Intermediate Results and Summary Statistics ........................................................................... 63
  5.2 Difference in Difference Regressions ........................................................................................ 70
  5.3 Durbin-Watson Tests for Autocorrelation ................................................................................ 76

Chapter 6: Discussion ............................................................................................................................ 80
  6.1 Trade Secret Litigation .............................................................................................................. 80
  6.2 Influence of the Empirical Method ............................................................................................ 81
  6.3 Summary of Central Findings .................................................................................................... 84
  6.4 Aggregate Behavior of the Control Coefficients ..................................................................... 85
  6.5 Dynamics of the Treatment Coefficients ................................................................................ 87
  6.6 Potential Negative Autocorrelation of Residuals ..................................................................... 89

Chapter 7: Conclusion ............................................................................................................................ 91
  7.1 Enhancing the Study Beyond the Scope of this Thesis ............................................................. 91
  7.2 Future Directions in Trade Secret Research ............................................................................ 93
List of Tables

3.1 Regressions Coefficients as a Function of Observation Characteristics for a Simple Difference in Difference Model............................................................44

3.2 Regressions Coefficients as a Function of Observation Characteristics for a Multi-Period Difference in Difference Model........................................52

5.1 Summary Statistics for Mean and Relative Effectiveness of Intellectual Property Protection Mechanisms Across Product and Process Innovations........64

5.2 Summary Statistics for Total Number of Patent Applications per State for Mapped Patent Classes in Pennsylvania and New Jersey for Seven Years Beginning March 15, 2001.................................................................68

5.3 Summary Statistics for Total Number of Patent Applications per State for Non-Eliminated Patent Classes in Pennsylvania and New Jersey for Seven Years Beginning March 15, 2001.................................................................69

5.4 Difference in Difference Regressions of Patent Application Time Series Data at the State and Composition of Industry Sector Levels........................73

5.5 Difference in Difference Regressions of Patent Application Time Series Data at the Industry Sector and Patent Class Levels........................................74

5.6 Difference in Difference Regressions of Trade Secret Litigation Time Series Data........................................................................................................75

5.7 Sign and Relative Magnitude of the Treatment Coefficients from Difference in Difference Regressions of Patent Application Time Series Data................76

5.8 Durbin-Watson Tests for First-Order Autocorrelation of the Residuals from Difference in Difference Regressions of Patent Application Time Series Data........................................................................................................78

5.9 Durbin-Watson Tests for First-Order Autocorrelation of the Residuals from Difference in Difference Regressions of Trade Secret Litigation Time Series Data.................................................................79

A1.1 Mean and Relative Effectiveness of Intellectual Property Protection Mechanisms Across Product and Process Innovations by Industry..............95

A1.2 Mean and Relative Effectiveness of Intellectual Property Protection Mechanisms Across Product and Process Innovations by Industry Sector........96

A1.3 Total Number of Patent Applications for Seven Years Beginning March 15, 2001 for Mapped Patent Classes in Pennsylvania and New Jersey.........................96


A1.5 Time Series Data for Trade Secret Litigation in Pennsylvania and New Jersey......................................................................................................100
List of Figures

1.1 Research Classifications by Motivation and Desired Outcome as Represented by Historical Practitioners ................................................................. 15
1.2 Disclosure Mechanisms for Codified Knowledge .......................................................................................................................... 18
1.3 Classification of Non-Disclosed Codified Knowledge .................................................................................................................. 20
5.1 Classifications of Industries into Industry Sectors .................................................................................................................. 65
A1.1 Surjective Mapping of OTAF Sequence Numbers onto Industry Sectors ....... 101
Chapter 1: Introduction to the Economics of Trade Secrecy

Chapter 1 provides context for the concept of trade secrecy. It begins broadly by discussing endogenous economic growth and the role technological innovation plays in such growth. This is followed by an explanation of how technological innovation occurs and how the knowledge from that innovation is disclosed. The chapter then discusses the conditions under which new knowledge is not disclosed and defines trade secrecy within that space. This discussion is followed by an introduction to economic theories of trade secrecy and the incentives therein. The chapter then incorporates trade secrecy into Romer's endogenous growth model. Chapter 1 concludes by assessing the importance of trade secrecy to the United States economy.

1.1 Endogenous Growth and Codified Knowledge

Understanding the dynamics of sustainable, long-term economic growth is one of the central research topics in macroeconomics. The importance of economic growth, whether in a nation’s gross domestic product (GDP), per capita income, or other indicator of economic prosperity, is grounded in the idea of multiplicative compounding. The power of compounding is best demonstrated by example. Imagine a lottery in which the sponsoring organization offers one of the following options as a prize:

1. A one-time lump sum payment of $1,000,000 today.

2. One penny today and sequential payments, once per day for one month, where each sequential payment is twice as much as the previous payment.

Which is more valuable? While one might be tempted to select option 1 and take the $1,000,000 today, option 2 is in fact worth $10,737,418;1 significantly more than the lump sum. Now imagine that the game show improved option 1 so that instead of receiving $1,000,000 today, the $1,000,000 would grow by 5% per day for one month, and the value

---

1 Under the assumption that one month is 30 days long and rounding to the nearest dollar.
at the end of the month would be the new prize. Even in this case, the prize would be less than half of the award in option 2, namely $4,116,136. The reason for this is in the difference in the growth rates. Even though the initial prize is 100,000,000 times less in option 2 than in the improved option 1, the fact that it grows 20 times faster makes it more valuable. While this example demonstrates that the rate of compounding, or growth, is significantly more important than the initial value to long run prosperity, sustainability of that growth is also critical. For instance, if option 2 had instead paid the contestant every other day for one month, the prize would come to only $328; an inconsequential amount when compared to the aforementioned prizes. While macroeconomic growth rates are often not as extreme as those used in the above example, their importance to the long-term prosperity of nations is clear nonetheless.

Endogenous growth theory, the most recent major contribution to the theoretical understanding of economic growth, was first put forth by Paul Romer and others in the late 1980s and early 1990s. This theory made several important modifications to Robert Solow’s neoclassical growth model, which introduced the idea of technological change as the cornerstone of economic growth. While Solow’s model was a major contribution to macroeconomics, earning him the 1987 Nobel Prize in Economics, it nonetheless allowed technological change to be determined exogenously. In addition, Solow’s model assumed diminishing returns to capital, which causes growth to slow as capital is increased. Romer and his contemporaries incorporated technological change into models of economic growth, and in doing so created endogenous growth theory. One of the hallmarks of

---


endogenous growth theory is that "knowledge is assumed to be an input in production that has increasing marginal productivity."\textsuperscript{6} This frees capital from diminishing returns as new codified knowledge, which is a product of human capital and existing knowledge, is in fact more useful to a society as it accumulates.\textsuperscript{7} In turn, growth is no longer slowed by the accumulation of capital, as long as the right balance of physical capital and new codified knowledge is being produced.\textsuperscript{8}

Codified knowledge is able to play such an important role in economic growth because of its unusual property of being non-rival and partially excludable.\textsuperscript{9} The concepts of rivalry and excludability can be applied in the analysis of any economic good and provide insight into how the good can be used in a group setting.\textsuperscript{10} Rivalry is an intrinsic property of a good and governs how many entities can use it simultaneously. A purely rival good can only be used by one agent at a time whereas a purely non-rival good can be used by a theoretically unlimited number of agents simultaneously. Excludability is determined both by the intrinsic properties of a good as well as the legal system that the good resides in, and represents whether an owner can prevent others from using the good. A purely excludable good is one in which the owner has complete control over its use, whereas a purely non-excludable good is one in which the owner has no control over its use. Taken together, the concepts of rivalry and excludability can be used to classify similar goods. Codified knowledge is non-rival because there is no \textit{a priori} limit to the number of agents that can use a given design, process, or philosophy simultaneously. However, codified knowledge is partially excludable in our society as the creators of that knowledge can exercise intellectual property rights, in the form patents, copyrights, and trademarks, which limit the use of that knowledge or even choose to keep the knowledge a secret.\textsuperscript{11}

Therefore codified knowledge not only acts cumulatively, due to its non-rival nature, but

\begin{itemize}
  \item\textsuperscript{6} Romer, "Increasing Returns and Long-Run Growth."
  \item\textsuperscript{7} Romer, "Endogenous Technological Change."
  \item\textsuperscript{8} Ibid.
  \item\textsuperscript{9} Ibid.
  \item\textsuperscript{10} Richard Cornes and Todd Sandler, \textit{The theory of externalities, public goods, and club goods} (Cambridge University Press, 1996).
\end{itemize}
can be used by profit seeking economic agents because of its partial excludability. Having established the importance economic growth and the critical role that the creation of new codified knowledge plays in sustaining that growth, it is now imperative to assess the mechanisms governing the creation of codified knowledge.

1.2 Creation of Codified Knowledge

Research, the activity governing the creation of codified knowledge, has two fundamental characteristics that determine its nature: motivation and desired outcome. Motivation for research can be broadly classified as either a fundamental curiosity or practical inspiration. Likewise, one can broadly classify desired outcomes of research as either furthering human understanding or solving a practical problem. Taken together, these classifications make a two by two matrix that can be used to classify similar research projects. See Figure 1.1 for a representation of this matrix. The quadrants in this matrix are named for scientists whose beliefs and work exemplified their various characteristics. For example, research motivated by fundamental curiosity but with the desired outcome of solving a practical problem is known as research in Pasteur’s quadrant. Understanding these classifications of research is important because researchers’ pursuing different quadrants will likely produce different kinds of codified knowledge. Moreover, different types of institutions often have differing motivations and goals and therefore pursue research in the quadrants that best match their incentives.

12 Donald E. Stokes, Pasteur’s quadrant: basic science and technological innovation (Brookings Institution Press, 1997).
The key to understanding the types of codified knowledge produced by our society is to analyze the incentives facing research performing organizations and the researchers that work in them. Research performing organizations can be further broken down into governments, for profit entities such as businesses, and non-profit organizations such as research universities and non-governmental organizations (NGOs). Due to the multifaceted, complex, and expansive role governments play in our society, they perform and or fund research in every quadrant. For example, the U.S. federal government allocated 12.8% of its discretionary budget for research in fiscal year 2008, approximately $138 billion. Nearly 60% that was set aside for the Department of Defense (DoD), which performs research almost exclusively in Bacon’s and Pasteur’s quadrants, with the remaining 40% divided

---

13 It is important to make the distinction between a research performing organization and a research organization. The former is any organization that performs research, including governments, universities, and firms. The latter, on the other hand, generally refers to an organization whose primary goal is research. While some research performing organizations are also research organizations, such as research universities, others, such as governments, are clearly not.

14 While this topic could be the focus of many theses on its own, my intention is to present a simple first-order analysis of the incentives surrounding these organizations. Any further depth is beyond the scope of this thesis.


among the National Science Foundation (NSF), National Institutes for Health (NIH), and other agencies.\textsuperscript{17} For profit entities are generally motivated to perform research by a desire to meet some unmet or partially met need. Furthermore, the end goal of their research is not merely to learn more about the need, but to create and market a product or service that meets it.\textsuperscript{18} This fits the paradigm of Baconian research, practically inspired and performed to solve real world problems. Unlike for profit institutions, non-profits have a wider range of motivations for research. While research universities have been the traditional home of research driven by fundamental curiosity with desire to simply increase human understanding, Newtonian research, this is not universally true. For example, the Massachusetts Institute of Technology (MIT) performs a significant amount of research in all three of the other quadrants through its School of Engineering and Lincoln Laboratory.\textsuperscript{19}

In addition, research universities get a significant portion of their funding from governments, aligning many of the research goals between these organizations. The motivation and goals for research in different organizations not only determine the type of research pursued, but how the results of that research are disclosed.

\section*{1.3 (Non)-Disclosure of Codified Knowledge}

Research performing organizations, such as for profit businesses and non-profit institutions (including research universities), can disclose the results of their research through two primary vehicles: publications and patents. However, these methods are not

\textsuperscript{17} National Science Foundation, "Federal R&D."

\textsuperscript{18} There are certainly exceptions to this paradigm, most notably Bell Laboratories under AT&T and Alcatel-Lucent. Despite being part of a for profit company, Bell Laboratories remains one of the world's pre-eminent research institutions having employed scientists and engineers who earned a combined 7 Nobel Prizes in Physics, 9 U.S. Medals of Science, 12 U.S. Medals of Technology, and many other prestigious awards while working at Bell Laboratories. See Alcatel-Lucent, "Awards and Recognition," Alcatel-Lucent: Bell Labs: About Bell Labs, http://www.alcatel-lucent.com/wps/portal/BellLabs/AwardsandRecognition#tabAnchor2 (Accessed on July 23, 2010).

\textsuperscript{19} The MIT Lincoln Laboratory "is a federally funded research and development center chartered to apply advanced technology to problems of national security. Research and development activities focus on long-term technology development as well as rapid system prototyping and demonstration. These efforts are aligned within key mission areas. The Laboratory works with industry to transition new concepts and technology for system development and deployment." See Lincoln Laboratory, "About Lincoln Laboratory," Home Page of the MIT Lincoln Laboratory, http://www.ll.mit.edu/about/about.html (accessed July 23, 2010).
mutually exclusive. As a result, one can create a two by two matrix that encompasses the vast majority of disclosure mechanisms for codified knowledge. See Figure 1.2 for a visual representation of this matrix. The upper right and lower left quadrants are the most familiar, publishing or patenting alone. Publications are the currency of academic study, the mechanism by which researchers gain both social recognition and future resources in the form of promotions and grants. While the acquisition of future resources through publication is more closely associated with research performed in a university setting, it is nonetheless true in some industries; most notably those associated with the biological sciences. Therefore both academic and industry researchers have a strong incentive to publish. On the other hand, firms generally prefer to disclose their discoveries in the form of patents. The reason for this is simple, patents provide firms with protection against competitors in the form of a 20 year, legally sanctioned monopoly. The government grants this monopoly in exchange for the disclosure of the invention so that others may take advantage of its non-rival nature. In doing so, the patent system aims to balance the commercial incentive to innovate with the necessary accumulation of codified knowledge to promote economic growth.

---

20 This is not to say that there are no restrictions on publication with regard to patenting. Specifically, one cannot patent an invention if it has been published or in public use for one year or more. As a result, agents almost always file a patent application (if applicable) before publishing their results. See 35 U.S.C. § 102(b).
As previously mentioned, patenting and publishing are not mutually exclusive, and when done together are known as a patent-paper pair.\textsuperscript{25} Essentially this is a dual-disclosure mechanism in which academic kudos are granted while protecting the commercial interests of the researcher’s employer. In fact, patent-paper pairs are used in both university and business settings.\textsuperscript{26} While the benefits of dual disclosure are clear, there are situations in which a research performing organization would opt to only pursue one of the two disclosure mechanisms. For example, a university researcher may not believe that his or her discovery has commercial value and therefore would forego the time, cost and effort of applying for a patent. On the other hand, a firm may believe that because a certain innovation is worth patenting, there is too much risk in disclosing a detailed description of the process leading up to that innovation for fear of losing some competitive process advantage. In fact, there are even reasons for research performing organizations to choose the lower right quadrant: non-disclosure.


\textsuperscript{26} Since the passing of the Bayh-Dole Act in 1980, universities have had the option to patent the inventions of their researchers. See 35 U.S.C. §§ 200-212 (2009).

---

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
\textbf{Decision to Patent} & \textbf{Patent} & \textbf{Not Patent} \\
\hline
\textbf{Decision to Publish} & & \\
\hline
Publish & Paten-Paper Pair & Publication \\
\hline
Not Publish & Patent & Non-Disclosure \\
\hline
\end{tabular}
\end{table}
While publishing, patenting, and patent-paper pairs have been and continue to be extensively studied, the non-disclosure of codified knowledge has been left by the proverbial wayside. However, results of research in this neglected fourth quadrant can be described by two criteria: usefulness of discovery and effort to not disclose. This creates a two by two matrix within the lower right quadrant of the matrix in Figure 1.2. A visual representation of this matrix can be found in Figure 1.3. The “usefulness of discovery” of a research finding is a measure of whether the result provides utility, either commercially or academically, to society. The term “effort to not disclose” is not a property of the finding itself, but of the researchers and or their research performing organization with respect to the finding. Specifically, it is a measure of the level of effort used to ensure that the finding is not disclosed. For example, one would expect to find a failed research project useful, but passively non-disclosed; placing it in the upper left quadrant. The knowledge that a specific research project failed would likely be useful to the research community at large, so that it could avoid wasting effort repeating the study and or improve upon they study to increase future chances of success. In addition, it is unlikely that the researchers who performed the failed study actively hid all evidence of their project, but more likely that they passively did not publish their failed study. If one moves down the matrix to non-useful, passively non-disclosed results, one finds that this lower left quadrant is occupied with common knowledge. Unlike the failed experiment, one that produces no new knowledge is not useful to the research community. Furthermore, there is no reason for the researcher’s who performed the study to actively hide the information, but it is unlikely they would attempt to publish it either. Moving over to non-useful, actively non-disclosed results, one finds that this lower right quadrant is populated by cases of mistakes and misconduct. Clearly findings that are the result of faulty research are not only not useful to the research community, but could even be detrimental if mistaken for sound research. Moreover, it is likely that researchers who realized they made an egregious error or committed misconduct would actively try and hide the results of that error or misconduct in an effort to preserve their reputation. This leaves us with the final and most interesting quadrant, the one containing useful and actively non-disclosed results. Namely, trade secrets.
### Figure 1.3 Classification of Non-Disclosed Codified Knowledge

<table>
<thead>
<tr>
<th>Effort to Not Disclose</th>
<th>Passive</th>
<th>Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Useful</td>
<td>Failed Research Project</td>
<td>Trade Secrecy</td>
</tr>
<tr>
<td>Not Useful</td>
<td>Common Knowledge</td>
<td>Faulty Research</td>
</tr>
</tbody>
</table>

### 1.4 Trade Secrecy Defined

The two criteria for knowledge to be in the upper right hand quadrant of Figure 1.3 are that the knowledge must be useful to society, in a commercial or academic context, and that the knowledge must be actively kept secret. This almost exactly mirrors the United States federal government’s definition of a trade secret:

all forms and types of financial, business, scientific, technical, economic, or engineering information, including patterns, plans, compilations, program devices, formulas, designs, prototypes, methods, techniques, processes, procedures, programs, or codes, whether tangible or intangible, and whether or how stored, compiled, or memorialized physically, electronically, graphically, photographically, or in writing if —

(A) the owner thereof has taken reasonable measures to keep such information secret; and

(B) the information derives independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable through proper means by, the public. \(^{27}\)

However, it is important to note the federal definition of a trade secret is slightly narrower than the criteria for being in the useful, actively non-disclosed quadrant in Figure 1.3.

Specifically, the federal government restricts the usefulness criterion to “economic value.”\textsuperscript{28} While this makes trade secrets a subset of the knowledge in the upper right quadrant of Figure 1.3, the matrices in subchapter 1.3 are no less useful at placing trade secrecy in the context of the disclosure of codified knowledge and economic growth.

1.5 Choosing Between Patent and Trade Secret Protection

Firms, the subset of commercial research performing organizations, often view trade secrecy as an alternative intellectual property protection mechanism to patenting. The reason for this is that trade secrets receive legal protection from “misappropriation,” which for the purposes of Chapter 1, can simply be thought of as theft or unauthorized disclosure.\textsuperscript{29} Specifically, firms that possess trade secrets can seek injunctive relief and damages from a misappropriator.\textsuperscript{30} However, firms that possess trade secrets can seek no relief if a competitor independently invents, reverse engineers, or observes in public use or display a trade secret and then uses that trade secret for economic benefit. Another important feature of trade secrecy is that trade secret protection never expires, so long as the knowledge remains secret and continues to provide economic benefit. In addition, any piece of codified knowledge can be protected as a trade secret as long as it meets the broad criteria outlined in Subchapter 1.4. However, patent protection has much more stringent criteria for protection.\textsuperscript{31} This implies that there is a nontrivial subset of innovations for which patent protection is not available and so trade secrecy is the default alternative. Although trade secrecy complements the patent system in this way, there are many more innovations for which firms have a choice of intellectual property protection mechanism. The above differences between trade secret and patent protection lead to several interesting theoretical conclusions about the types of innovations and firms that use make use of each form of protection when presented with a choice.

\textsuperscript{28} Ibid.
\textsuperscript{29} Subchapter 2.1 presents more detailed definitions of “misappropriation” throughout the history of trade secret law in the United States.
Theory predicts that firms more frequently protect complex and intermediate innovations as trade secrets and its patents. Complex innovations are generally very hard to reverse engineer, making reinvention the only non-misappropriation-based method for a competitor to acquire a complex innovation protected by trade secrecy. In essence, trade secret protection acts as a patent with an indefinite lifetime for complex innovations. Interestingly, even if a complex innovation is reinvented, the reinventing competitor often has no way of verifying that they have in fact re-created the product or process. The reason for this is that it is often just as difficult to determine if two mixtures or alloys are equivalent entities to determine their composition. This means that even if the product is reinvented, the original trade secret holder can continue to market his or her product as if it were unique.

Intermediate innovations, however, are not necessarily difficult to reverse engineer. The reasons firms prefer to protect intermediate innovations through trade secrecy are that patenting them is time-consuming, costly, and can provide information to competitors about the final product giving the competitors an opportunity to invent around the intermediate patent and beat the original firm to market with the final product. The value of an intermediate innovation is determined by its ability to quickly and inexpensively move the inventing firm closer to marketable product. In addition, intermediate innovations lose their value after the creation of the final marketable product unless they can serve as intermediate innovations for another product. Therefore taking the time and money to patent them both directly and negatively impacts their value. In addition, the

---

32 A complex innovation is a product or process that is composed of many different non-easily separable components. For example, many chemicals in food products are considered complex whereas a biotechnology therapeutic, which is essentially a purified protein, is not. Intermediate innovations are those that will eventually compose a final product, but are not marketed themselves. Many process innovations are considered intermediate.


innovating firm gains little from the patent protection as product cycles are often shorter than the years it takes for the United States Patent and Trademark Office (USPTO) to approve the patent. However, patent applications are published only 18 months after they received. This could potentially give competitive firms, especially those that have more efficient development organizations, to see the intermediate patent and use it to develop and patent a final product that could then exclude the original innovating firm's product. The timescales involved in the US patent system are also relevant in determining whether certain types of firms prefer to use trade secrecy than patent protection.

Firms with innovation cycles that differ significantly from the lifetime of patent protection may favor trade secrecy. Firms whose innovation cycles are significantly shorter than 20 years may prefer trade secrecy to patent protection because in the time it takes for a patent to be approved on a product, the next generation of that product may be available. This significantly reduces the value of the patent on the first product. On the other hand, firms whose innovation cycles are significantly longer than 20 years may prefer trade secrecy to patent protection because they are guaranteed to lose their competitive advantage in 20 years with a patent, whereas a trade secret and give them a chance at perpetuating this advantage. In addition, long innovation cycles imply costly research and development. This means that the information contained in the patent is significantly more valuable to competitive firms who could save research and development costs by using that knowledge to develop a different product outside of the patent's claims. This means that by patenting, the innovating firm would be subsidizing their competitors' research. While this is always true, it matters significantly more in industries that have relatively high costs of research and development. However, innovation cycle length is not the only firm characteristic that impacts the decision of whether to use trade secrecy or patent protection.

---


Firm size also influences firms’ choice of intellectual property protection. Specifically, small firms often rely much more heavily on trade secrecy than large firms. The reason for this is that patents are only valuable if the patent holder has the resources to launch an infringement suit. Often, small and entrepreneurial firms lack these resources and could not sustain an extended litigation against the multinational corporation. Therefore by patenting, these firms would simply be describing their innovations to the world with little ability to defend them. While it is also necessary to engage in a legal dispute to be compensated for trade secret misappropriation, protecting innovations through trade secrecy removes the issue of publicly revealing the fact that the innovation even exists. While this does little to help protect product innovations that are easily reverse engineered, it does provide a viable solution for processing innovations.

1.6 Endogenous Growth and Trade Secrecy

It is clear from the discussion in Subchapter 1.5 that there are many situations in which a firm would prefer to use trade secrecy than patent protection for a given innovation. However, one of the core assumptions in Romer’s model of endogenous technical growth is that “anyone engaged in research has free access to the entire stock of knowledge. This is feasible because knowledge is a non-rival input. All researchers can take advantage of at the same time.” While this assumption may have been useful to simplify the model to make more understandable and to study its basic dynamics, a technique used later in this thesis to introduce difference in difference analysis, it is patently false given the existence of trade secrecy. To understand the effects of trade secrecy on economic growth, Romer’s model must be modified.

---

39 Horstmann, MacDonald, and Slivinski, “Patents as Information Transfer Mechanisms.”
40 is the total stock of “designs” in the economy, in other words, the complete set of codified knowledge. See Romer, “Endogenous Technological Change.”
The first section of Romer's model that changes with the introduction of trade secrecy is his calculation of the rate of production of new codified knowledge, or designs in his terminology, by each researcher. Specifically, Romer states that if the researcher possesses an amount of human capital $H_i$ and has access to a portion $A_i$ of the total stock of knowledge implicit in previous designs, the rate of production of new designs by researchers will be $\delta H_i A_i$, where $\delta$ is a productivity parameter.

Given the assumption stated in the previous paragraph, Romer then states that this production rate is equivalent to $\delta H_i A$ because $A$ is equal to $A_i$ if each researcher has access to the entire body of codified knowledge. To introduce trade secrecy, let $s$ be the fraction of innovations kept secret in the economy. In other words, $s$ the percent of innovations that are protected as trade secrets instead of patents. Then, $A_i = A(1 - s) + A_{private}$

where $A(1 - s)$ is the codified knowledge known by every researcher and $A_{private}$ is the secret knowledge known by the individual researcher. Since the sum of each individual researcher’s private knowledge must be the total amount of knowledge in the economy less the common knowledge, $A = A(1 - s) + \sum_{j=1}^{N} A_{private}^j$

$As = \sum_{j=1}^{N} A_{private}^j$

where $N$ is the total number of researchers in the economy. Romer then finds the growth rate of codified knowledge in the economy.

If we sum across all people engaged in research, the aggregate stock of designs evolves according to

41 It is important to clarify that in Romer's model, researchers are the holders of patents and creators of innovations. Therefore, he examines the economy at the level of the researcher. Translating this into reality, where firms are the holders of patents and creators of innovations, one simply replaces the word "researcher" with firm.

42 Romer, "Endogenous Technological Change."
\[ \dot{X} = \delta H_A A \]

where \( H_A \) has the obvious interpretation of total human capital employed in research.\(^43\)

This summation is slightly more complicated with the secrecy parameter \( s \) and is shown below.

\[
\dot{A} = \sum_{j=1}^{N} \delta H_j^j (A(1-s) + A_{private}^j)
\]

\[
\dot{A} = \delta A (1-s) \sum_{j=1}^{N} H_j^j + \delta \sum_{j=1}^{N} H_j A_{private}^j
\]

\[
\dot{A} = \delta H_A A (1-s) + \delta \sum_{j=1}^{N} H_j A_{private}^j
\]

\[\delta H_A A \geq \dot{A} \geq \delta H_A A (1-s)\]

where the growth rate of codified knowledge and the economy depends on the distribution of human capital and private knowledge among firms. The left-hand side of the fourth equation above corresponds to the case where all of the human capital and private knowledge and the economy are held by one researcher, whereas the right-hand side corresponds to an economy in which no researcher possesses a nonzero amount of human capital and private knowledge. As both of these extremes bear little resemblance to reality, the assumption that each researcher possesses an identical amount of human capital or secret knowledge will be made.\(^44\) This implies

\[
\dot{A} = \delta H_A A (1-s) + \frac{\delta H_A A s}{N}
\]

\[
\dot{A} = \delta H_A A (1-s) \frac{(N-1)}{N}
\]

\(^{43}\) Ibid.

\(^{44}\) While this assumption is not significantly more realistic than the previous two, as the distribution of firm size is thought to be exponential, it is a common economic assumption that is useful in understanding the model. The reason this statement uses the "or" conjunction is both of the forms have the same result.
As the number of in an economy is often quite large, this equation can be considered functionally equal to

\[ \dot{A} = \delta H_A (1 - s) \]

for the purposes of this investigation. Romer continues his analysis and arrives at the growth rate

\[ g = \delta H_A \]

Carrying this factor of \((1-s)\) through Romer’s analysis, one arrives at the modified growth rate

\[ g = \delta H_A (1 - s) \]

This implies that according to our current understanding of economic growth is driven by endogenous technological change, the growth decreases linearly in the frequency of trade secrecy. While this is more of an estimation that result, and its conclusion needs to be thoroughly empirically tested, it does provide a potentially serious concern about the economic benefit of trade secret protection.

1.7 Importance of Trade Secrecy in the U.S. Economy

While economic theory promotes our understanding of trade secrecy when it occurs, theory provides little insight into the relative frequency of trade secrets and patents. Unfortunately for academics, there is no comparable database to the patent records by the USPTO for trade secrets. The reason for this is obvious, as any such database of trade secrets would negate all value in the secrets therein. This makes empirical studies of trade secrecy difficult and accounts for why there are so few. Two primary methods are employed by researchers when investigating questions relating to trade secrecy. Specifically, these methods are the use of industry-wide surveys and the use of litigation

\[ \text{Romer, } "\text{Endogenous Technological Change."} \]
records to search for trade secret cases. While both of these methods are indirect, as answers to surveys are notoriously imprecise and the relationship between the number of trade secrets and trade secret cases is unknown, they do provide a tool to investigate aggregate phenomena where precise results are less important than general trends.

The most striking result from empirical on trade secrecy research is that trade secrets are viewed by firms as significantly more effective at protecting both product and process innovations than patents in almost all industries.\textsuperscript{46} Interestingly, as theory predicts, this effect decreases as firm size increases.\textsuperscript{47} The most cited reasons for using trade secrecy instead of patent protection are reported, in descending order, as the ease of inventing around patents, not being able to meet the "novelty" requirement for patent protection, the amount of information disclosed in the patent, and the application costs.\textsuperscript{48} This seems to imply that a significant percent of innovation in firms is rooted in inventing around existing patents. If firms were instead creating novel innovations, lack of novelty would not be ranked above amount of information disclosed. This evidence implies, but certainly does not confirm, that the patent system is failing to provide existing firms with protection over their ideas, but instead serves as a way for young firms and inventors to secure rights over their creations.

If one is to believe that firms in the vast majority of industries view trade secret protection as more effective than patent protection, then the effect of trade secrecy on the United States economy is as powerful as that of the patent system. Yet, our understanding of trade secrets and their effect on the economy is negligible in comparison to our understanding of the patent system and its effects. Therefore empirical research on trade secrecy presents one of the largest and most important unexplored research topics in economics.

\textsuperscript{46} Cohen, Nelson, and Walsh, "Protecting Their Intellectual Assets."
\textsuperscript{48} Cohen, Nelson, and Walsh, "Protecting Their Intellectual Assets."
Chapter 2: Trade Secret Law

Chapter 1 introduced the concept of trade secrecy in the context of economic growth and codified knowledge as well as engaged the current economic understanding of trade secrecy. Chapter 2 provides a legal framework from which to understand trade secrecy. It begins by reviewing the history of trade secret law in the United States and in turn how the current definition of trade secrecy, given in Subchapter 1.4, developed. The chapter then presents contemporary trade secret law as a form of statutory intellectual property law. Chapter 2 concludes by revealing that the research community knows very little about how the use of statutory trade secret law affects the economics of trade secrecy.

2.1 History of Trade Secret Law in the United States

Legal protection of trade secrecy is older than any formal conception of intellectual property. While there is some debate over the existence of explicit trade secret protection in the Roman Empire, trade secret law certainly dates back to the trade guilds of Renaissance Europe.49 The laws that protected guilds from the unauthorized use of their secret knowledge during the Renaissance developed into schemes for rewarding for inventors and limiting the mobility of skilled labor in the early modern period.50 Despite the fact that trade secrets predate formal conceptions of intellectual property and have been a feature of European laws for centuries, the United States Constitution made no reference to trade secrecy even though the legal bases for patents and copyrights were

---


included. This relegated trade secrecy to common law protection in the United States, which began to take form in the mid-nineteenth century in Massachusetts.

The first known case relating to trade secrecy in the United States was in herd by the Supreme Court of Massachusetts in 1837 and set the precedent for courts to recognize the misappropriation of trade secrets as cause for damages. In 1866, the same court considered injunctive relief against the misappropriation of a trade secret in Taylor v. Blanchard. However, the court dissolved the non-compete agreement that would have provided that relief because “although [the process] was not generally known to the public, it was carried on in three different towns in the Commonwealth, by three different parties, who had no connection in business with the parties to this contract;” thus formally introducing the condition of “secrecy” into trade secret law. Building from these piecemeal rulings, the first comprehensive judicial opinion on trade secrecy was given just two years later. In 1868, the Supreme Court of Massachusetts heard Peabody v. Norfolk in which Francis Peabody sued John Norfolk for an injunction against a factory Norfolk was building to manufacture gunny cloth from jute butts using a technique he had learned while employed by Peabody under a non-disclosure and non-compete agreement. The court found for Peabody, stating:

It is the policy of the law, for the advantage of the public, to encourage and protect invention and commercial enterprise. If a man establishes a business and makes it valuable by his skill.

---

51 U.S. Const. art. I, § 8, cl. 8.
53 In Vickery v. Welch, 36 Mass. (19 Pick.) 523, 527 (1837), the plaintiff sued the defendant for breaching the conditions of a contract by failing to give the plaintiff exclusive rights to the defendant’s secret manner of making chocolate after the plaintiff had purchased the defendant’s chocolate business. The court held that by not restricting himself to secrecy, the defendant had breached the contract and would be held liable for damages.
54 In Taylor v. Blanchard, 95 Mass. (13 Allen) 370 (1866), the plaintiff was a shoe-cutter manufacturer who partnered with the defendant and in doing so entered into a non-compete agreement covering the state of Massachusetts. After the partnership “dissolved by mutual consent,” the defendant opened his own shoe-cutter manufacturing business in a different city. The plaintiff sued the defendant for breach of contract and appealed a negative decision to the Massachusetts Supreme Court, but the court held that the agreement was void and saw “nothing beneficial to the public” in the non-compete agreement as it would either promote a monopoly or draw workers away from the state. The fact that the court found that the confidential information was not in fact secret, as there were three other manufacturers in the Commonwealth, was crucial to the reasoning.
57 Peabody v. Norfolk (1868).
and attention, the good will of that business is recognized by the law as property. If he adopts and publicly uses a trade mark, he has a remedy, either at law or in equity, against those who undertake to use it without his permission. If he makes a new and useful invention of any machine or composition of matter, he may, upon filing in a public office a description which will enable an expert to understand and manufacture it, and thus affording to all persons the means of ultimately availing themselves of it, obtain letters patent from the government securing to him its exclusive use and profit for a term of years. If he invents or discovers, and keeps secret, a process of manufacture, whether a proper subject for a patent or not, he has not indeed an exclusive right to it as against the public, or against those who in good faith acquire knowledge of it; but he has a property in it, which a court of chancery will protect against one who in violation of contract and breach of confidence undertakes to apply it to his own use, or to disclose it to third persons. The jurisdiction in equity to interfere by injunction to prevent such a breach of trust, when the injury would be irreparable and the remedy at law inadequate, is well established by authority. 58

There are several key ideas presented in the above passage that formed the basis of trade secret law in the United States. First, the statement that "it is the policy of the law, for the advantage of the public, to encourage and protect invention and commercial enterprise" implies that in cases of trade secrecy, and intellectual property in general, it is the court's duty to ensure incentives to invent are maintained. 59 Second, the court explicitly states that a breach of trade secrecy would qualify for damages and possibly injunctive relief. 60 Finally, the above paragraph clearly places trade secrecy in the domain of intellectual property by comparing it to trademarks and patents as well as labeling it "property." 61

However, this notion was altered by the United States Supreme Court in 1917 in the opinion of E.I. du Pont de Nemours Powder Co. v. Masland which viewed the "misappropriation of trade secrets ... as a tort based on the confidential relationship between the parties or the misbehavior of the defendant." 62 The idea of trade secret law as tort seemingly replaced the conception of it as intellectual property and was prominently featured in the first attempt to consolidate trade secret law.

58 Ibid.
59 Ibid.
60 Ibid.
61 Ibid.
62 E.I. du Pont de Nemours Powder Co. v. Masland, 244 U.S. 100, 102 (1917); In E.I. du Pont de Nemours Powder Co. v. Masland, "the plaintiff corporation sought certiorari review of a judgment of the United States Court of Appeals for the Third Circuit, which reversed a decree of the district court and dissolved an injunction prohibiting defendant ex-employee from disclosing certain alleged trade secrets of plaintiffs to experts retained in preparation for the defense of plaintiffs' suit to prevent defendant from using the trade secrets." See E.I. du Pont de Nemours Powder Co. v. Masland; Lemley, "Surprising Virtues of Treating Trade Secrets as IP Rights, The."
The Restatement (First) of Torts was published by the American Law Institute in 1939 and included a section entitled "Misappropriation of Trade Secrets" in an attempt to "unify the nascent law of trade secrets." Since trade secret law had been organically developing through the common law, and mostly in the Commonwealth of Massachusetts, there was a significant need for a more universal definition of what constituted a trade secret and what recourse the possessor of one could have against misappropriation. The Restatement defines a trade secret as "any formula, pattern, device, or compilation of information which is used in one's business, and which gives him an opportunity to obtain an advantage over competitors who do not know or use it." Furthermore, it specifies that a substantial element of secrecy must exist, so that, except by the use of improper means, there would be difficulty in acquiring the information. An exact definition of a trade secret is not possible. Some factors to be considered in determining whether given information is one's trade secret are:

1) the extent to which the information is known outside of his business;
2) the extent to which it is known by employees and others involved in his business;
3) the extent of measures taken by him to guard the secrecy of the information;
4) the value of the information to him and to his competitors;
5) the amount of effort or money expended by him in developing the information;
6) the ease or difficulty with which the information could be properly acquired or duplicated by others.

The Restatement also specifies conditions under which acquisition of a trade secret can be considered misappropriation:

One who discloses or uses another's trade secret, without a privilege to do so, is liable to the other if

(a) he discovered the secret by improper means, or
(b) his disclosure or use constitutes a breach of confidence reposed in him by the other in disclosing the secret to him, or
(c) he learned the secret from a third person with notice of the facts that it was a secret and that the third person discovered it by improper means or that the third person's disclosure of it was otherwise a breach of his duty to the other, or
(d) he learned the secret with notice of the facts that it was a secret and that its disclosure was made to him by mistake.

Almeling, "Four Reasons to Enact a Federal Trade Secrets Act;" Restatement (First) of Torts §§ 757-758 (1939).

Restatement (First) of Torts § 757 (1939).

Ibid.

Ibid.
In addition, the Restatement clarifies that a third party who learns of a trade secret without knowing that it is a trade secret is not liable until he or she becomes aware that the knowledge is a trade secret.\(^\text{67}\) Despite widespread acceptance of the Restatement, courts were able to adopt it in a piecemeal fashion as it was not binding.\(^\text{68}\) This was compounded by the fact that trade secret law continued to develop unevenly, “as states in commercial centers developed extensive case law while agricultural states had a leaner body of precedent.”\(^\text{69}\) As a result, the Restatement largely failed to bring full uniformity to trade secret law. As a response to this continued heterogeneity, the National Conference of Commissioners on Uniform State Laws proposed a statutory Uniform Trade Secrets Act (UTSA) in 1968 for adoption by the states as an alternative to their continue reliance on common law.\(^\text{70}\) However, there were several notable federal statutes and U.S. Supreme Court cases that shaped trade secret law in the time between the Restatement (First) of Torts and the UTSA’s initial adoption in 1979.

In the mid-twentieth century, the United States federal government began to recognize trade secrecy as an institution. In 1948, the federal government outlawed the distribution of trade secrets that federal agencies may become privy to their work by those agencies or employees of those agencies.\(^\text{71}\) This provision was of vital importance as the FDA and other regulatory agencies were coming into increasing contact with sensitive corporate information, often in the form of trade secrets. When the Freedom of Information Act (FOIA) was introduced in 1966, it contained (and still contains) a provision that exempts trade secrets from request.\(^\text{72}\) In addition, the Occupational Health and Safety Act (OHSA) or 1970 includes a provision that protects the confidentiality of trade secrets.\(^\text{73}\) In fact, the OHSA protection is so strong that if there were a disaster of the magnitude of


\(^{68}\) Almeling, “Four Reasons to Enact a Federal Trade Secrets Act;”

\(^{69}\) Almeling, “Four Reasons to Enact a Federal Trade Secrets Act.”

\(^{70}\) Uniform Trade Secrets Act (1985).


Bhopal on U.S. soil and the harmful agent was a trade secret, only medical personnel and others directly involved in the disaster management process would be informed of the name and properties of the agent. The federal government's recognition of trade secrecy by way of these additions to the United States Code was supplemented by Federal and Supreme Court rulings in the 1970s and 1980s.

In 1970, the United States Court of Appeals for the 5th Circuit heard *E. I. du Pont de Nemours & Co. v. Christopher* in which it held that "improper means" as outlined in the Restatement did not mean illegal, but instead below "standard of morality expected in our commercial relations." Additionally, the court cited its role in preserving "the spirit of inventiveness" by protecting firms from spending exorbitant sums on protecting their secrets. These justifications for trade secrecy law continue to the present day. Just four years later in 1974, the United States Supreme Court heard *Kewanee Oil Co. v. Bicron Corp.*, which established that neither the Patent Clause of the United States Constitution nor federal patent laws preempt state trade secret protection for patentable or unpatentable codified knowledge. In addition, the court stated that "the patent policy of encouraging invention is not disturbed by the existence of another form of incentive to invention. In this respect the two systems are not and never would be in conflict." While the court's assumption that patents and trade secrets are purely complementary is a clear fallacy, it

---

74 This is not to say that public pressure would not either 1) force the company responsible for the disaster to disclose the chemical willingly or 2) prompt an executive order overriding the OHSA. However, as written, the public would not be informed of the agent without an intervention of the kind mentioned above.

75 In *E. I. du Pont de Nemours & Co. v. Christopher*, 431 F.2d 10122 (5th Cir. 1970), the defendants were hired to fly over the construction site of the plaintiff's new plant to take aerial photographs of the design. The plaintiff filed suit alleging that the defendants had improperly obtained photographs revealing a trade secret and had sold those photographs to a third party. After several motions and appeals, the United States Court of Appeals for the 5th Circuit found in favor of the plaintiff.


77 *Kewanee Oil Co. v. Bicron Corp.*, 416 U.S. 470 (1974) began when Harshaw Chemical Co. brought a case to the United States District Court for the Northern District of Ohio seeking injunctive relief and damages against former employees, who were then employees of Bicron Corp., for the misappropriation of trade secrets relating to a crystal growing technology. The District Court's decision to grant a permanent injunction was appealed to the Court of Appeals for the Sixth Circuit which overturned the ruling finding that Ohio trade secret law was in conflict with United States patent law in that Ohio could not grant monopoly protection to processes and manufacturing techniques which were eligible for patenting. The Supreme Court of the United States granted certiorari to resolve the question of the preemption of state trade secret law by federal patent law.

78 *Kewanee Oil Co. v. Bicron Corp.*
serves as the foundation for placing trade secrets on equal legal footing with patents as intellectual property protection mechanisms. This decision also swung the definitional pendulum of trade secret law back from being thought of as a tort to being considered a form of intellectual property. Despite the marked increase in the power of trade secret protection as a result of *Kewanee*, the most widely used definition of trade secrecy was revoked just four years later when the American Law Institute omitted the section entitled "Misappropriation of Trade Secrets" from the Restatement (Second) of Torts. Fortunately, the National Conference of Commissioners on Uniform State Laws had already drafted several versions of the UTSA by 1978, and the omission by the American Law Institute gave them reason to act.

On August 9, 1979 the National Conference of Commissioners on Uniform State Laws approved the UTSA and recommended it be enacted in all states. Subsequently amended in 1985, the UTSA was the first statutory trade secret law in the United States and attempted to resolve the heterogeneity in state common law through voluntary unification. At its core, the UTSA is simply a set of definitions for terms and conditions for relief that state legislatures to enact if they so choose. The UTSA defines the terms improper means, misappropriation, person, and trade secret as

1. "Improper means" includes theft, bribery, misrepresentation, breach or inducement of a breach of a duty to maintain secrecy, or espionage through electronic or other means;

2. "Misappropriation" means:
   (i) acquisition of a trade secret of another by a person who knows or has reason to know that the trade secret was acquired by improper means; or
   (ii) disclosure or use of a trade secret of another without express or implied consent by a person who
      (A) used improper means to acquire knowledge of the trade secret; or
      (B) at the time of disclosure or use, knew or had reason to know that his knowledge of the trade secret was
         (I) derived from or through a person who had utilized improper means to acquire it;

---

79 Restatement (Second) of Torts § 757 (1978).
81 Ibid.
82 It is important to recognize that the UTSA was created by a nongovernmental organization and so was not binding in any jurisdiction upon initial approval.
(II) acquired under circumstances giving rise to a duty to maintain its secrecy or limit its use; or
(III) derived from or through a person who owed a duty to the person seeking relief to maintain its secrecy or limit its use; or
(C) before a material change of his [or her] position, knew or had reason to know that it was a trade secret and that knowledge of it had been acquired by accident or mistake.

(3) "Person" means a natural person, corporation, business trust, estate, trust, partnership, association, joint venture, government, governmental subdivision or agency, or any other legal or commercial entity.

(4) "Trade secret" means information, including a formula, pattern, compilation, program, device, method, technique, or process, that:
   (i) derives independent economic value, actual or potential, from not being generally known to, and not being readily ascertainable by proper means by, other persons who can obtain economic value from its disclosure or use, and
   (ii) is the subject of efforts that are reasonable under the circumstances to maintain its secrecy.83

In addition, the UTSA explicitly provides courts with the following rules for awarding injunctive relief and damages:

SECTION 2. INJUNCTIVE RELIEF.

(a) Actual or threatened misappropriation may be enjoined. Upon application to the court, an injunction shall be terminated when the trade secret has ceased to exist, but the injunction may be continued for an additional reasonable period of time in order to eliminate commercial advantage that otherwise would be derived from the misappropriation.

(b) If the court determines that it would be unreasonable to prohibit future use in exceptional circumstances, an injunction may condition future use upon payment of a reasonable royalty for no longer than the period of time the for which use could have been prohibited. Exceptional circumstances include, but are not limited to, a material and prejudicial change of position prior to acquiring knowledge or reason to know of misappropriation that renders a prohibitive injunction inequitable.

(c) In appropriate circumstances, affirmative acts to protect a trade secret may be compelled by court order.

...
SECTION 3. DAMAGES.

(a) In addition to or in lieu of injunctive relief, except to the extent that a material and prejudicial change of position prior to acquiring knowledge or reason to know of misappropriation renders a monetary recovery inequitable, a complainant may be entitled to recover damages for the actual loss caused by misappropriation. A complainant also may recover for damages caused by misappropriation and the unjust enrichment caused by misappropriation that is not taken into account in computing damages for actual loss. In lieu of damages measured by any other methods, the damages caused by misappropriation may be measured by imposition of liability for a reasonable royalty for a misappropriator's unauthorized disclosure or use of a trade secret.

(b) If willful and malicious misappropriation exists, the court may award exemplary damages in an amount not exceeding twice any award made under subsection (a). 84

Notice that these definitions are significantly more detailed than those in the Restatement (First) of Torts. The reason for this is that while both documents sought to "unify" trade secret law, the Restatement was written as a set of principles and common definitions to guide judicial opinions, whereas the UTSA was written to serve as an explicit set of rules for courts to implement. The idea behind the UTSA is that if a state adopts the Act, the courts of that state would have little to no flexibility in its implementation. In fact, Section 8 of the UTSA goes so far as to say that the Act "shall be applied and construed to effectuate its general purpose to make uniform the law with respect to the subject of this Act among states enacting it." 85

To date, 46 states have adopted the UTSA. 86 Given the wording of the Act discussed in the previous paragraph, this might lead one to believe that it achieved its goal of uniformity. However, this could not be farther from the truth. While the vast majority of states adopted the Act, many did so in a piecemeal fashion, altering the definition of a trade secret. 84

---

84 Ibid. Text that has been struck through with included in 1979, but not in the amended version. Underlined text was added in 1985 amendment.
85 Uniform Trade Secrets Act (1985). The four that currently rely exclusively on the common law or trade secret protection are New York, New Jersey, Massachusetts and Texas. It is interesting to note that the two states with the strongest trade secret common law, New York and Massachusetts, are among the non-adopters.
86 Almeling, "Four Reasons to Enact a Federal Trade Secrets Act."
secret, what constitutes misappropriation and even the conditions under which injunctive relief and damages can be awarded.\(^{87}\) In addition, many states did not adopt Section 8 and even in those that did, courts often ignored the mandate and offered minority opinions.\(^{88}\) Even at the federal level, specifically in United States district courts, versions of the UTSA are only cited as the applicable law in between 50% and 80% of cases per year.\(^{89}\) However, trade secret litigation rates in district courts have grown exponentially since 1980.\(^{90}\) While no study has yet shown causation between the UTSA and this exponential growth, the correlation certainly merits investigation.

In the 1990’s, the federal government began to play a more active role in trade secret law for the first time in the history of the United States. For example, the Economic Espionage Act of 1996 makes the theft or misappropriation of trade secrets by a foreign entity a federal crime.\(^{91}\) While the Economic Espionage Act does provide a definition for the term “trade secret,” as presented in Subchapter 1.4, it does little to unify the law beyond criminal trade secret law because the definition is not applicable outside the statute and is largely similar to the one provided by the UTSA. While unification was not the purpose of the statute, it is important to recognize that a federal trade secrets act could easily achieve what the Restatement (First) of Torts and the UTSA failed to if it were proposed and enacted. The federal government also passed the American Inventor’s Protection Act of 1999, which applies the concept of prior user rights to trade secrets. Specifically, it protects trade secret holders from being sued for patent infringement for the use of their long held trade secrets. The passing of the American Inventor’s Protection Act is particularly notable for two reasons. First, these prior user rights significantly increase firms’ incentive to protect their intellectual property through trade secrecy instead of through the patent system by removing one of the two major risks associated with trade secrecy.\(^{92}\) Second, it

\(^{87}\) Ibid.

\(^{88}\) Ibid.


\(^{92}\) The other, more important, risk is the loss of competitive advantage.
represents the first time the federal government has introduced a concept into the cannon of trade secret law.\textsuperscript{93} This sets the precedent for federal involvement in the continued development of statutory trade secret law and lays the foundation for a potential federal trade secrets act.

2.2 Trade Secret Law as Statutory Intellectual Property Law

The two most salient features of the history of trade secret law in the United States are that the legal basis of the law is poorly defined and that over the past three decades, trade secret law has begun to transition from being primarily a feature of the common law to being statutorily defined. By virtue of being left out of the United States Constitution and allowed to develop organically through the common law, courts were able to decide for themselves whether trade secret law was fundamentally a tort, a contract, a form of intellectual property, some combination of these, or a branch of law all its own. While contemporary legal scholars continue to debate the origin and proper legal designation of trade secret law using both normative and historical arguments, this debate has little affect on how trade secrets are used in the United States economy.\textsuperscript{94} \textit{Kewanee Oil Co. v. Bicron Corp.} reestablished the original doctrine of viewing trade secret law as a form of intellectual property in the United States, a principle the federal government statutorily reinforced with the American Inventor’s Protection Act of 1999. While Supreme Court decisions can be reversed, the fact that the federal government supported that classification, albeit implicitly, with statutory law is a strong signal of the stability of this paradigm. However, the federal government’s enactment of statutory trade secret law is more relevant to the future of trade secret protection when viewed in the context of the ongoing common law to statutory transition than as a confirmation that in practice, trade secret law operates as a form of intellectual property law.

\textsuperscript{93} The only other major federal involvement in statutory trade secret law, the Economic Espionage Act of 1996, did not fundamentally change any characteristics of trade secret law. It simply adopted the definition of the UTSA, with some very minor modifications, and applied it in a new way.

From the time of the Constitution to the late 1970s, trade secrets had been exclusively protected under the common law of each individual state. While the Restatement (First) of Torts attempted to guide the development of state common law toward more uniform standards, the UTSA was the first attempt at a statutory trade secret law in the United States. Since the UTSA's creation in 1979, United States trade secret law has largely shifted from the common law to statutory law. However, heterogeneity among state laws has remained. By involving itself in the development of statutory trade secret law with the passage of the American Inventor's Protection Act of 1999, the federal government has taken the first step toward formalizing trade secret law as statutory intellectual property law. It is important to understand that if the federal government did formalize trade secret law as statutory intellectual property law, it would also implicitly be unifying trade secret law across the United States. This then leads to the question, should the United States have a federal trade secrets act?

While this question is difficult to assess as phrased above, it can be more easily understood when divided into two sub-questions:

1. Are the economic effects of trade secrecy aligned with the interests of the United States?

2. How does statutory law influence the use of trade secrecy?

If the answers to both of these questions are known, then the federal government can design a statutory trade secret law to alter the use of trade secrecy in a way that promotes the interests of our society. If the answer to the first question was known, but not to the second, then the federal government would know how it wanted to influence trade secret use through a federal trade secrets act, but not how to craft legislation to foster that influence. If the answer to the second question is known, but not to the first, then the federal government would understand the influence of their legislative choices on trade secret use, but would not know what the desired result of a federal trade secrets act should be. If the answer to neither question is known, then the federal government would neither know how it wanted to influence trade secret use in the United States nor how to craft
legislation to implement a desired influence. Currently, this is the state of understanding. Chapter 1 engaged the positive component of the first question in an effort to understand the economic effects of trade secrecy and observed that while there are numerous theories, they are supported by little to no empirical evidence. Despite the significant amount of research attempting to answer the positive component of the first question, there has been no economic work on the second. This is surprising given the shift in trade secret protection from the common law to statutory law over the last several decades. This thesis intends to fill this gap by exploring how the implementation of statutory trade secret law affects the selection of intellectual property protection and subsequently, the disclosure of codified knowledge.

95 The first question can be divided into positive and normative components. The positive component is the question "what are the economic effects of trade secrecy?" The normative component is the question "what are the interests of our society?" If both of these questions are answered, then the answer to the first question follows. In addition, the normative component of the first question cannot be answered by economic or legal research as the interests of the United States are determined exogenously.

96 This is not to say that arguments have not been made for, and against, the implementation of a federal trade secrets act, but that there have been no economic studies exploring the direct effect of statutory law on trade secret use. However, there have been many legal studies that explore the standards used by courts before and after legislation including, the UTSA. See Almeling, David S. et al, "A Statistical Analysis of Trade Secret Litigation in Federal Courts."
Chapter 3: Empirical Design

Chapter 2 introduced the concept of trade secret law and outlined the uncertainty associated with the effects of statutory trade secret law. Chapter 3 describes an empirical approach designed to gain insight into those uncertain effects. It begins by reframing the Uniform Trade Secrets Act (UTSA) as a natural experiment and introducing difference in difference analysis as an empirical tool to measure the aforementioned uncertain effects. The chapter then evaluates possible states and observables for use in the analysis and discusses the relevant experimental parameters. This is followed by an adaptation of the basic difference in difference model to the aforementioned states, observables, and parameters. Chapter 3 concludes by describing the expected effects of an implementation of the UTSA and in turn the expected results of the analysis in Chapter 4.

3.1 The Uniform Trade Secrets Act: A Natural Experiment

The incomplete and heterogeneous adoption of the UTSA across the United States provides a natural experiment for studying the effects of implementing statutory trade secret law on knowledge disclosure. In Subchapter 2.1 I introduced the UTSA as the most recent attempt to unify United States trade secret law at the state level. In addition, I observed that the UTSA largely failed to unify trade secret law in the United States due to its incomplete and heterogeneous adoption. However, this policy failure created an opportunity for the research community to examine how knowledge disclosure is influenced by the implementation of statutory trade secret law through the comparison of similar states with varying levels of UTSA adoption. Natural experiments that exhibit this form, namely comparable groups of which some undergo a change at a specific point in time and others do not, lend themselves to difference in difference analysis.
Difference in difference analysis is an empirical tool that isolates the effect of a treatment on the treated group. In this case, the treatment is the adoption of the UTSA in a given state and the group is the set of firms, organizations, and individuals in that state who face knowledge disclosure decisions. As the goal is to assess how the treatment affects the treated, i.e. how the adoption of the UTSA affected the knowledge disclosure decisions of the aforementioned entities, an ideal metric would be the number of new trade secrets created each year. While this is practically impossible, as there are no publically accessible databases of trade secrets, imagine for the moment that one could measure the number of new trade secrets per year as easily as one could measure the number of new patents per year. If this were the case, one might be tempted to simply calculate the measured difference in the rate of trade secret creation before and after the adoption of the UTSA as the treatment’s effect on the treated. However, that method is unable to distinguish between the effect of the treatment and any time trends that may also be affecting the treated group. On the other hand, one might be tempted to use the set of firms, organizations, and individuals who face knowledge disclosure decisions in an untreated state with similar time trends as a control and calculate the effect of the treatment on the treated as the difference in the number of trade secrets between the treated state and the control state after the treatment. However, this method is unable to distinguish between the treatment effect and a difference in the initial endowment, i.e. the number of trade secrets created by the states before the treatment. By combining these two methods, namely taking the difference of the difference between each state’s pre-treatment and post-treatment trade secret creation rates, one can remove both biases simultaneously.

The simplest difference in difference model has a single treatment group, a single control group, and divides time into two periods: pre and post treatment. Regressions of this model take the form

$$y_{it} = \beta_0 + (s_i)(\beta_2) + (t)(\beta_3) + (s_i)(t)(\beta_4) + \epsilon_{it}$$

98 It is important to note that databases of trade secrets do exist. Notably, New Jersey maintains the New Jersey Trade Secret Registry to track trade secrets in the chemicals industry. See 34 N.J.S.A § 5A-15 (2009).
where \( y_i \) is the \( i \)th observation, \( s_i \) is an indicator variable for the treated state, \( t \) is an indicator variable for the second time period (treatment), \( \beta_j \) is the \( j \)th regression coefficient, and \( \varepsilon_i \) is the error associated with the \( i \)th observation.\(^99\) By examining Table 3.1, one can see that \( \beta_1 \) represents the baseline level of observation in the control state, \( \beta_2 \) represents the difference between the baseline level of observation in the treatment state and the control state,\(^100\) \( \beta_3 \) represents the general time trend in the data\(^101\) and \( \beta_4 \) represents the difference in difference estimation of the effect of the treatment on the treated.\(^102\) While this simple form is useful for understanding the fundamental idea behind difference in difference analysis, a more complex difference in difference model is needed to examine the impact of implementing statutory trade secret law on knowledge disclosure. In order to derive that model, I must first identify the observational variables and states I intend to use in the analysis.

### Table 3.1 Regression Coefficients as a Function of Observation Characteristics for a Simple Difference in Difference Model

<table>
<thead>
<tr>
<th>State ((s_i))</th>
<th>Time Period ((t))</th>
<th>Regression Coefficients ((\beta_j))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s_i = 0)</td>
<td>(t = 0)</td>
<td>(\beta_1)</td>
</tr>
<tr>
<td>(s_i = 0)</td>
<td>(t = 1)</td>
<td>(\beta_1 + \beta_3)</td>
</tr>
<tr>
<td>(s_i = 1)</td>
<td>(t = 0)</td>
<td>(\beta_1 + \beta_2)</td>
</tr>
<tr>
<td>(s_i = 1)</td>
<td>(t = 1)</td>
<td>(\beta_1 + \beta_2 + \beta_3 + \beta_4)</td>
</tr>
</tbody>
</table>

### 3.2 Observing Trade Secret Litigation and Patent Applications

While direct measurement of trade secrets can be used in a thought experiment like the one discussed in the previous subchapter, there is currently no practical way to acquire direct time series data on trade secrets as of 2010. As discussed in Subchapter 1.7,

\[^{99}\text{In the example from the previous paragraph, the observational variable in Equation 3.1 would be the average number of new trade secrets created in the associated time period and state. Additionally, the second time period would begin when the UTSA was enacted in the treated state.}\]

\[^{100}\text{\((\beta_1 + \beta_3) - \beta_1 = \beta_2\). This controls for the difference in initial endowments.}\]

\[^{101}\text{\((\beta_1 + \beta_3) - \beta_1 = \beta_3\). This controls for the existence of a general time trend in the data.}\]

\[^{102}\text{\(((\beta_1 + \beta_2 + \beta_3 + \beta_4) - (\beta_1 + \beta_2)) - ((\beta_1 + \beta_3) - \beta_1)\) = \beta_4.}\]
researchers who study trade secrecy are beginning to use time series data on trade secret litigation as an indirect method of empirical measurement. However, this technique is complicated by the fact that trade secret litigation rates are a function of both the existing number of trade secrets and of the probability that a given trade secret will be involved in litigation. For example, if trade secret litigation increased in response to some policy shock, then either the number of trade secrets increased, the propensity for firms to litigate increased, or (more likely) both increased. While separating these effects may seem impossible, I contend that their relative impact can be estimated. Any significant and immediate change in the number of trade secret cases in response to a policy shock should be exclusively associated with a change in the propensity to litigate. The reason for this is that no firm could have had time to create a trade secret, either from a new development or previously unprotected intellectual property, have had the secret misappropriated, and then pursued litigation. Therefore the change could not be due to an increase in the underlying number of trade secrets. In addition, the propensity to litigate would not be expected to change over time after a policy shock because it is able to attain a new equilibrium value instantaneously. Yet, one would expect the number of litigations to change over time after a policy shock, relative to a control, as either trade secrets that would have been patents or publications accumulate or patents and publications are created that would have been trade secrets. Therefore one could estimate the initial change in response to a policy shock as a shift in the propensity to litigate while subsequent variation over time could be understood as changes in the underlying number of trade secrets. While the implementation of this estimation technique further supports the use of trade secret litigation time series data to study trade secrecy, examining other mechanisms of knowledge disclosure is similarly important to understanding the effect of the implementation of statutory trade secret law on knowledge disclosure.

103 Almeling et al., “A Statistical Analysis of Trade Secret Litigation in Federal Courts;”
104 I recognize that this is a simplification and that an alternative argument could be made that there is no immediate change in propensity to litigate because firms wait to see how the courts interpret the new statute and then change their propensity to litigate over time in response to that interpretation. However, it is clear that the propensity to litigate not only changes on a faster time scale than the underlying number of trade secrets, but also takes significantly less time to be observed through trade secret litigation. Therefore I believe this simplification is a useful estimation tool.
While patents are often the focus of research on knowledge disclosure, patent applications are far more informative to researchers studying potential changes in the composition of aggregate knowledge disclosure. When an entity is deciding whether to patent, protect its IP through trade secrecy, or follow some other knowledge disclosure route such as publication, the decision is being made at the patent application stage. The reason for this is clear: organizations that would prefer a patent to a trade secret file a patent application, whereas those that would prefer trade secret protection do not file an application. This distinction was reinforced on March 15, 2001 when the United States Patent and Trademark Office (USPTO) began to publish patent applications, eliminating the possibility of protecting the contents of a rejected application through trade secrecy.105 The patent approval process therefore acts as a confounding factor when trying to examine knowledge disclosure decisions because not all patents are approved. In addition, the patent approval process creates a lag between intent to patent, the variable of interest when examining changes in knowledge disclosure incentives, and the granting of a patent. For all of the above reasons, shifts in knowledge disclosure behavior from patenting to trade secrecy, or vice versa, are best observed by examining time series data on patent applications and trade secret litigation.

By examining changes in the levels of trade secrecy litigation and patent applications in response to the implementation of a statutory trade secret law, such as the UTSA, one can assess the effect of the law on aggregate knowledge disclosure. One can reasonably expect that a change in trade secret law would affect entities deciding whether to use patent or trade secret protection and entities deciding whether to keep something a trade secret or use an alternative disclosure mechanism, but not entities deciding between patenting and alternative disclosure mechanisms such as publication. Therefore one can assume that any change in the number of patent applications caused by the implementation of the UTSA is a result of organizations choosing to use trade secrecy more, or less, at the margin. Due to the poorly defined structure of alternative mechanisms of

105 Interestingly, this transformed patent applications from compliments of trade secrets, as one could apply for a patent without revealing the secret and withdraw the application before approval and publication if one so chose, to substitutes. However, patents themselves have always been substitutes for the subset of patentable trade secrets because of the inherent disclosure.
knowledge disclosure, such as non-academic publication, there are no comparable observables to trade secret litigation or patent applications for these disclosure mechanisms. Therefore determining the effect of the implementation of the UTSA on organizations that frequently choose between trade secrecy and these mechanisms must be done implicitly, and therefore is more difficult a less precise. For example, if patent applications and the component of trade secret litigation that, according to the estimation in the previous paragraph, corresponds to the underlying change in the number of trade secrets both increase or decrease as a result of the implementation of the UTSA, then the effect on knowledge disclosure on the border between trade secrecy and alternative disclosure mechanisms is unclear. In addition to examining the effect of the implementation of the UTSA on aggregate knowledge disclosure, it is also interesting to examine the effects on individual industries, industry sectors and patent classes. Once the observables have been selected, in this case trade secret litigation and patent applications on a statewide basis in aggregate as well as on individual industries, industry sectors and patent classes, the next step is to establish the specific treatment selected for study.

3.3 The Pennsylvania Uniform Trade Secrets Act

In this analysis I will examine the effect of the implementation of the Pennsylvania Uniform Trade Secrets Act (PAUTSA) on knowledge disclosure in the state of Pennsylvania. The PAUTSA was first considered by the Pennsylvania legislature on February 11, 2003, signed into law on February 19, 2004, and took effect two months later on April 19, 2004. Not only is Pennsylvania the most recent state to adopt a version of the UTSA, but it is also the only state to do so since the USPTO began publishing patent applications on March 15, 2001. This makes the adoption of the PAUTSA the only UTSA implementation for which time series data can be collected on patent applications before and after the adoption. Specifically, I will collect 10 years of trade secret litigation data beginning on April 19, 1999 as well as seven years of patent application data beginning on March 15, 2001.106

2001 and use the treatment dates of April 19, 2004 and March 15, 2003 respectively.\textsuperscript{107} The treatment date for trade secrecy litigation was simply chosen as the effective date for the PAUTSA, as it was the first day that the law could be recognized in court. The treatment date for patent applications was chosen to reflect the ability of organizations to anticipate changes in legislation when the bill is being considered while also accommodating the availability of data.\textsuperscript{108} Beyond the ability to directly observe the effect on patent applications, the PAUTSA was adopted with only minor changes to the UTSA, making it an ideal candidate for gaining insight into the potential effects of a national trade secret act.\textsuperscript{109} In fact, the one alteration was to permit criminal penalties for trade secret misappropriation, a characteristic that the PAUTSA shares with the federal definition of trade secrecy in the Economic Espionage Act of 1996.\textsuperscript{110} However, no matter how intriguing a particular policy shock may be, it cannot be analyzed using difference in difference regressions without a suitable control group.

Fortunately, trade secrecy litigation and patent applications in New Jersey form a suitable control group. New Jersey is similar to Pennsylvania in a macroeconomic sense and is one of the four remaining states not to have adopted a version of the UTSA. In fact, difference in difference comparisons using New Jersey and Pennsylvania are well established in the literature.\textsuperscript{111} Therefore I intend to use patent application in trade secret

\textsuperscript{107} Since the USPTO publishes patent applications 18 months after the effective filing date, the years beginning March 15, 2008 at March 15, 2009 at incomplete patent application data and therefore were left out this study. While I would have preferred to collect 10 years of data on patent applications, five years before and five years after treatment, the ability to directly observe patent applications outweighs the advantage an extra three years pretreatment data would bring to the study because it would require studying patents instead of applications.

\textsuperscript{108} Ideally I would have used February 11, 2003 as the treatment date for patent applications because it is the first time an organization could have anticipated the enactment of the PAUTSA. However, patent applications became available on March 15 two years prior. Therefore for consistency I chose March 15 as the treatment date. In fact, the slightly less conservative treatment date of March 15, 2003 may even provide more accurate results.

\textsuperscript{109} Since the UTSA was created with the intention of uniform state adoption, it was in essence an attempt to create a national trade secrets act at the state level, an analysis of the effects of its adoption in a state that modified the act significantly when adopting it would provide less general insight because of the existence of clauses specific to that state.


litigation time series data from New Jersey as the control. While there has been a trend in recent years to use synthetic control states in difference in difference analyses, the fact that trade secret litigation is best observed on the state level in combination with this study's use of a single treatment state limits the pre-treatment observations to a level that precludes the creation of a synthetic control state. Now that the observables, states, and experimental parameters have been introduced, I can adapt the basic difference in difference model to analyze the effect of the adoption of the PAUTSA on knowledge disclosure in Pennsylvania.

3.4 Adaptation of the Basic Difference in Difference Model

A multi-period difference in difference model can be used to analyze the impact of implementing statutory trade secret law, such as the PAUTSA, on knowledge disclosure strategies. To generalize the model presented in Subchapter 3.2 to multiple periods, one must first decouple the treatment dummy from the time and state dummies. Without decoupling these dummies, it would be impossible to create a model with multiple pre-treatment time periods. The second step in generalizing the two period model is to modify the time indicator so that it captures the dynamics of the time series data. To accomplish this, one could simply include time dummies for each period. Regressions using this generalized model take of the form

\[ y_i = \beta + (s_i)\beta_2 + (f_i)\beta_3 + (s_i)(f_i)\beta_4 + \left( \sum_{n=1}^{N} t_{n} (\beta_{4n+1} + (s_i)\beta_{4n+2} + (f_i)\beta_{4n+3} + (s_i)(f_i)\beta_{4n+4}) \right) + \epsilon_i \]

where \( y_i \) is the \( i \)th observation, \( s_i \) is an indicator variable for the treated state, \( f_i \) is an indicator variable for the treatment, \( t_n \) is an indicator variable for the \( n \)th time period, \( \beta_j \) is

---


113 In the two period model, the treatment dummy is implicitly included in the cross-term of the state and time dummies.
the $j^{th}$ regression coefficient, $N$ is the number of time periods after the initial time period, and $\epsilon_i$ is the error associated with the $i^{th}$ observation. By examining Table 3-2, one can see that $\beta_1$ represents the baseline level of observation in the control state, $\beta_2$ represents the difference between the baseline level of observation in the treatment state and the control state, $\beta_3$ represents the immediate effect of the treatment on the control, $\beta_4$ represents the difference in difference estimation of the immediate effect of the treatment on the treated, $\beta_{4n+1}$ represents the effect of $n$ time periods on observables in the control state, $\beta_{4n+2}$ represents the difference in difference estimate of the effect of $n$ time periods on observables in the treatment state, $\beta_{4n+3}$ represents the difference in difference estimate of the effect of the treatment after $n$ time periods on observables in the control state, $\beta_{4n+4}$ represents the difference in difference estimate of the effect of the treatment on the treated after $n$ time periods. However, this model is too general to be used in any difference in difference regression because it includes regression coefficients that are assumed to be zero by the difference in difference experimental design. Specifically, $\beta_3$ and $\beta_{4n+3}$ are assumed to be zero because the treatment should not affect observables in the control state and $\beta_{4n+2}$ is assumed to be zero because the control state is supposed to control for all non-treatment based time effects. While one might be tempted to include them for completeness, they interact with the error term and impair the regressions ability to accurately model the data. Taking these terms to zero, the new model becomes

$$
114 (\beta_1 + \beta_2) - \beta_1 = \beta_2. \text{This controls for the difference in initial endowments.}
$$

$$
115 (\beta_1 + \beta_2) - \beta_1 = \beta_3. \text{This is assumed to be zero in difference in difference regressions as the treatment should have no immediate effect on the observables in the control state.}
$$

$$
116 ((\beta_1 + \beta_2 + \beta_3 + \beta_4) - (\beta_1 + \beta_2)) - ((\beta_1 + \beta_3) - \beta_1) = \beta_4. \text{This is the immediate effect component of the overall difference in difference effect of the treatment on the control state.}
$$

$$
117 (\beta_1 + \beta_{4n+1}) - \beta_1 = \beta_{4n+2}. \text{This controls for time trends.}
$$

$$
118 ((\beta_1 + \beta_{4n+1} + \beta_2 + \beta_{4n+2}) - (\beta_1 + \beta_{4n+1})) - ((\beta_1 + \beta_2) - \beta_1) = \beta_{4n+3}. \text{This is assumed to be zero in difference in difference regressions as the (pure) effect of time on the observables in the control state and in the treatment state should be the same. While this may not be the case in practice, it is an assumption of a perfect control.}
$$

$$
119 ((\beta_1 + \beta_{4n+1} + \beta_3 + \beta_{4n+3}) - (\beta_1 + \beta_{4n+1})) - ((\beta_1 + \beta_3) - \beta_1) = \beta_{4n+4}. \text{This is assumed to be zero in difference in difference regressions as observables in the control state should not be affected by the treatment.}
$$

$$
120 ((\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_{4n+1} + \beta_{4n+2} + \beta_{4n+3} + \beta_{4n+4}) - (\beta_1 + \beta_2 + \beta_{4n+1} + \beta_{4n+2}) - (\beta_1 + \beta_{4n+1})) - ((\beta_1 + \beta_2 + \beta_3 + \beta_4) - (\beta_1 + \beta_2)) - ((\beta_1 + \beta_3) - \beta_1) = \beta_{4n+5}. \text{This is the time effect component of the overall difference in difference effect of the treatment on the treated.}
$$

$$
121 \text{Theoretically, in a scenario with a very large number of observations, and by consequence a very small error term, one could include these regression coefficients to check the accuracy of the control selection.}
$$
\[ y_i = \beta_1 + (s_i)\beta_2 + (s_i)(f_i)\beta_3 + \left( \sum_{n=1}^{N} t_n (\beta_{2n+2} + (s_i)(f_i)\beta_{2n+3}) \right) + \epsilon_i \]

However to use the above model in a regression on real data, one would need \(2N + 3\) observations to simply constrain the regression and significantly more to acquire results of any significance. Since trade secret litigation is best observed on the state level and I am only examining cases in Pennsylvania and New Jersey, the number of observations is \(2N\). In fact, there are not even enough observations in each set of patent application observations to support a regression of the form shown in Equation 3.3.\(^{122}\) Therefore I must make the simplifying assumption that the effect of the PAUTSA on the number of trade secret cases and patent applications per year in Pennsylvania is linear in time.\(^{123}\) Mathematically, this sets the each of the \(\beta_{2n+2}\) terms equal to some \(\beta_4\) and each of the \(\beta_{2n+3}\) equal to some \(\beta_5\).\(^{124}\) Therefore this final model, which is the one that is used for regressions in this study, has the form

\[ y_i = \beta_1 + (s_i)\beta_2 + (s_i)(f_i)\beta_3 + (t)(s_i)(f_i)\beta_5 + \epsilon_i \]

where \(t\) is equal to the time period of the observation. Note that \(t\) is not an indicator variable, making the cumulative effect of the treatment on the treated at time period \(t\) equal to \(\beta_3 + (t)\beta_5\). Now that the general structure of the difference in difference model that will be used in this study has been developed, is important to understand the ways in which its application to trade secret litigation and patent application time series data differs.

\(^{122}\) For example, the number of observations in each patent class is identical to the number of observations for trade secret litigation for the same reason. Specifically, that while one could subdivide a patent class into its subclasses or trade secret litigation by type of misappropriation, there would not be enough data in each of the subdivisions to overcome the noise. With approximately 500,000 patent applications per year, approximately half of which are of foreign origin, 50 states and approximately 300 active patent classes, the estimated average number of patent applications per class per state per year is fewer than 20. See United States Patent and Trademark Office, “U.S. Classes by Number with Title Menu,” Guidance, Tools, and Manuals, Classification, http://www.uspto.gov/web/patents/classification/selectnumwithtitle.html (accessed on August 8, 2010); United States Patent and Trademark Office, “U.S. Patent Statistics Chart: Calendar Years 1963-2009,” Patent Technology Monitoring Team, http://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm (accessed on August 10, 2010).

\(^{123}\) Is important to note that the model seen in Equation 3.2 does not assume a linear treatment effect in time. Since the \(\beta_{2n+1}\) terms are independent, the differences between them are not functionally constrained. Therefore they are able to detect nonlinear effects in time.

\(^{124}\) This is not the \(\beta_t\) from Equation 3.2, but a new regression coefficient.
Table 3.2 Regression Coefficients as a Function of Observation Characteristics for a Multi-Period Difference in Difference Model

<table>
<thead>
<tr>
<th>State ((s_i))</th>
<th>Time Period ((t_n))</th>
<th>Treatment ((f))</th>
<th>Regression Coefficients ((\beta_j))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(s_i = 0)</td>
<td>(t_n = 0)</td>
<td>(f_i = 0)</td>
<td>(\beta_1)</td>
</tr>
<tr>
<td>(s_i = 0)</td>
<td>(t_n = 0)</td>
<td>(f_i = 1)</td>
<td>(\beta_1 + \beta_3)</td>
</tr>
<tr>
<td>(s_i = 1)</td>
<td>(t_n = 0)</td>
<td>(f_i = 0)</td>
<td>(\beta_1 + \beta_2)</td>
</tr>
<tr>
<td>(s_i = 1)</td>
<td>(t_n = 0)</td>
<td>(f_i = 1)</td>
<td>(\beta_1 + \beta_2 + \beta_3 + \beta_4)</td>
</tr>
<tr>
<td>(s_i = 0)</td>
<td>(t_n = 1)</td>
<td>(f_i = 0)</td>
<td>(\beta_1 + \beta_4n+1)</td>
</tr>
<tr>
<td>(s_i = 0)</td>
<td>(t_n = 1)</td>
<td>(f_i = 1)</td>
<td>(\beta_1 + \beta_3 + \beta_4n+1 + \beta_4n+3)</td>
</tr>
<tr>
<td>(s_i = 1)</td>
<td>(t_n = 1)</td>
<td>(f_i = 0)</td>
<td>(\beta_1 + \beta_2 + \beta_4n+1 + \beta_4n+2)</td>
</tr>
<tr>
<td>(s_i = 1)</td>
<td>(t_n = 1)</td>
<td>(f_i = 1)</td>
<td>(\beta_1 + \beta_2 + \beta_3 + \beta_4 + \beta_4n+1 + \beta_4n+2 + \beta_4n+3 + \beta_4n+4)</td>
</tr>
</tbody>
</table>

The coefficients of the difference in difference model that will be used in this study have fundamentally different meanings for the trade secret litigation and patent application time series data that I proposed to observe in Subchapter 3.2. The reason for this is that the trade secret litigation data is a measure of the aggregate codified knowledge, the number of trade secrets, whereas the patent application data is a measure of the rate of change in aggregate codified knowledge per year.\(^{125}\) Therefore, for patent application data, Equation 3.4 represents an increasing marginal production of codified knowledge because \(\beta_4\) and \(\beta_5\) are the growth rates of the rate of change in aggregate codified knowledge per year. Specifically, Equation 3.4 represents a quadratic functional form for codified knowledge in time because the observation is modeled as a linear function and is akin to the first derivative of aggregate codified knowledge. On the other hand, for trade secret litigation, Equation 3.4 represents constant marginal production of codified knowledge because \(\beta_4\) and \(\beta_5\) are the rates of change in aggregate codified knowledge themselves. This is a linear functional form for codified knowledge in time. While the quadratic functional form used for patent application data is likely to better match the nonlinear growth model proposed by Romer, there is no observable equivalent for trade secrecy.\(^{126}\) The result is

\(^{125}\) It is important to understand that this is not a distinction between observing the number of new patent applications and the number of new patents, but the distinction between observing the number of patents versus observing the number of new patents. Therefore observing the number of new patents, instead of the number of patent applications, each year would also be a measure of the change in the aggregate amount of codified knowledge. However, for the reasons outlined in subchapter 3.2, observing the number of new patent applications per year is preferable to observing the number of new patents per year.

\(^{126}\) Quadratic functions are often used as an approximation for exponential functions, especially for short time periods. For a description of the nonlinear models mentioned herein, see Romer, "Endogenous Technological Change"; Romer, "Increasing Returns and Long-Run Growth."
that while one can still use the results of the trade secret litigation regression and the patent application regressions in concert to assess a change in the composition of codified knowledge, one cannot directly compare the regression coefficients. However, due to the fact that there is little evidence, if any at all, that allows the calculation of the number of trade secrets from the number of trade secret cases in any given jurisdiction, it would not be fruitful to compare the regression coefficients even if the models were identical. While difference in difference analysis is suited to testing the effect of the adoption of the PAUTSA on the number of trade secret cases and patent applications in Pennsylvania, one must be vigilant of the possibility of autocorrelation in the error terms.

Positively autocorrelated residuals have been shown to cause statistically significant overestimation of significance levels in many difference in difference studies.\textsuperscript{127} While this phenomenon is less common in studies with fewer control and treatment groups, such as this, it is nonetheless necessary to test for autocorrelation in the error terms and implement a correction if positive autocorrelation is detected.\textsuperscript{128} Specifically, I will use the Durbin-Watson statistic to test for first-order autocorrelation in the residuals of Equation 3.4. If positive autocorrelation is detected, I will be forced to further simplify the difference in difference analysis to a regression of the form of Equation 3.1 by averaging the pre-treatment and post-treatment time series data into two time periods. The reason for this potential simplification is that none of the other accepted correction methods is effective for studies with a small number of groups. Under the assumption that positive autocorrelation of residuals does not oversimplify this study, I will now discuss the expected results of the difference in difference analysis.


\textsuperscript{128} While one should correct for negative autocorrelation as well, it does not have the effect of overestimating significance levels. In fact, it has the opposite effect; namely, negative autocorrelation of the residuals can lead to serious underestimation of significance levels. Therefore it would be in a researcher's best interest performance correction nonetheless. See Ibid.
3.5 Expected Results

As discussed in Subchapters 1.8 and 2.4, the effect of implementing statutory trade secret law on knowledge disclosure is largely unknown. Logical arguments supporting a wide range of outcomes can be made a priori, especially since empirical research on the matter is scant at best. For example, an argument that supports an increase in trade secret use is that explicitly defined trade secret rights in a statutory trade secret law would likely promote the use of trade secrets because such rights remove the inherent uncertainty in common law decisions. On the other hand, an argument in support of no change in trade secret use is that the common law is so well established that explicit definitions of trade secret rights are irrelevant because they are ignored by the courts. Finally, an argument in support of a decrease in trade secret use is that the common law protects trade secrets in such a broad manner that a statute which explicitly defined reasonable trade secret rights would certainly reduce the broad protection currently granted by the courts. Despite the plethora of possible theories, there are logical clues that point to some theories over others. For example, the fact that 46 out of 50 states have implemented some version of the Uniform Trade Secrets Act (UTSA) implies that there is some need for an explicit definition of trade secret rights; however this in itself does not indicate whether the adopting states are attempting to promote or curtail trade secret use. However, the fact that two of the states with the strongest common law on trade secrets, namely Massachusetts and New York, are among the non-adopters, is an indicator that the UTSA is meant to enhance property rights and in turn promote trade secret use. Based on that assumption, I will now discuss the expected results of the analysis.

Since the direct output of the regressions will be a set of best-fit coefficients, I will organize the expectations around those coefficients beginning with the coefficients that are not part of the difference in difference estimate of the treatment on the treated. $\beta_1$, the baseline level of observation in the control state, is the least interesting parameter for both trade secret litigation and patent application rates. I simply expect $\beta_1$ to be positive, as the average number of either trade secret cases or patent applications in New Jersey cannot be
less than zero and is likely greater than zero. Furthermore, I expect \( \beta_1 \) to be close to the aforementioned average. \( \beta_2 \), the difference between the baseline level of observation in the treatment state and the control state, is slightly more interesting than \( \beta_1 \). While I am uncertain whether \( \beta_2 \) will be positive or negative and by how much, this uncertainty simply exists because there is no \textit{a priori} way of knowing whether Pennsylvania or New Jersey had more trade secret cases or patent applications in the first time period without first collecting the data on that time period. This will quickly be resolved during data collection. I expect \( \beta_3 \), the time trend of the observables in the control state, to be large and positive for both trade secret litigation and patent application rates. The rate at which knowledge has been codified has been increasing over the last several decades, and the 2000s were no exception. Therefore I expect this to be reflected in baseline growth of both trade secrets, as represented by trade secret litigation, and patent applications over time. I now turn my attention to the components of the difference in difference effect of the treatment on the treated.

\( \beta_3 \), the difference in difference estimate of the immediate effect of the treatment on the treated, is the first of the two proper results. For trade secret litigation, I expect \( \beta_3 \) to be positive because the propensity to litigate should increase as a result of the establishment of well-defined trade secret rights in an environment in which they are lacking.\textsuperscript{129} However, I have no prediction for its magnitude as the level of trade secrecy in our economy, which is by definition the upper bound, is unknown. For patent applications, I expect \( \beta_3 \) to be small compared to the untreated level at the time of treatment, \((\beta_1 + \beta_2 + 2\beta_4)\), and negative because an increase in trade secret protection should result in a substitution effect. However, this effect would likely be small compared to the aggregate number of patent applications because the strength of the statute in court is initially unknown. Therefore it is unlikely that firms would abandon patenting altogether in favor of uncertain trade secret protection. In addition, I expect \( \beta_3 \) to decrease in magnitude as one moves from industries,

\textsuperscript{129} However, if the alternative explanation given in supra 103 is to be believed, specifically that after a change from common law to statutory law firms wait to see how the courts will interpret the new statute before they bring new cases, then I would expect \( \beta_3 \) to be large and negative. The reason for this is that the propensity to litigate would drop sharply as firms held back potential cases waiting to see how the courts would interpret the new statute.
industry sectors and patent classes for which trade secrecy is relatively less important to those in which it plays a greater role because the potential substitution effect would be greatest felt by firms with trade secret to patent effectiveness ratios near one.\textsuperscript{130}

$\beta_5$, the difference in difference estimate of the additional effect of the treatment on the treated in each time period, is the second of the two proper results. For trade secret litigation, I expect $\beta_5$ to be positive because greater trade secret protection should both encourage research in areas that lead to codified knowledge that is traditionally protected through trade secrecy as well as create a substitution effect from patent applications to trade secrets. In addition, the fact that both knowledge creation and the process leading from the establishment of a trade secret to litigation about that secret take a non-negligible amount of time, there is strong reason to believe that the number of trade secret cases would continue to grow over time in response to a strengthening of trade secret protection. However, the relative magnitude of this growth cannot be predicted $a\ priori$. For patent application rates, I expect $\beta_5$ to be small compared to the untreated growth rate of patent applications, $\beta_4$, and negative. The expected sign of $\beta_5$ is purely due to the aforementioned substitution effect. Furthermore, the substitution is effect is very unlikely to result in a negative or nonexistent growth rate for patent applications because that would imply that trade secrecy cannibalized all of the increasing marginal production associated with the accumulation of new codified knowledge over time. Therefore the magnitude of the effect is likely to be small relative to the untreated growth rate of patent applications. Another result of the organizational variation is that the effect should be weaker in patent classes, industries and industry sectors for which trade secrecy is relatively important; in those with trade secrecy dependent effectiveness ratios close to one.\textsuperscript{131} Having introduced and discussed the empirical design, I will now delineate the exact method I used to implement it.

\textsuperscript{130} Recall that firms associated with a high relative importance of trade secrecy have an average trade secret patent effectiveness ratio of 2.24 whereas those associated with the low relative importance of trade secrecy have an average trade secret to patent effectiveness ratio of 1.32. Therefore firms with a low relative importance of trade secrecy are those for which patenting and trade secrecy are closer to being equally effective.

\textsuperscript{131} Ibid.
Chapter 4: Empirical Methodology

Chapter 3 described the study's empirical design. Chapter 4 delineates the exact method used to carry out that design so that fellow researchers can replicate the results if they so choose. It begins by describing how the industry sectors used in the study were selected and how they were refined according to patent classification. The chapter then discusses the data acquisition and classification procedure used in this study. Chapter 4 concludes by describing the adaptation of standard difference in difference regressions to patent application and trade secrecy litigation time series data.

4.1 Sector Selection and Refinement

The first step in implementing the research design outlined in Chapter 3 was to identify industries for which the relative importance of trade secrecy was high, average and low as compared to patenting. I began by examining Tables I and II in Cohen et al., which compare the effectiveness of appropriability mechanisms across different industries for product and process innovations.\textsuperscript{132} Using the columns corresponding to secrecy and patents, I created three additional metrics. To first two were the effectiveness of secrecy and the effectiveness of patenting across product and process innovations for each industry.\textsuperscript{133} The third metric was the ratio of the overall effectiveness of secrecy (the first new metric) to the overall effectiveness of patenting (the second new metric). I used this ratio of the effectiveness of secrecy to the effectiveness of patenting as a signal of the importance of trade secrecy in each industry and sorted the industries by this metric. I then clustered industries likely to patent in the same classes that also had similar ratios into industry sectors. I selected three sectors that contained industries with relatively high ratios, three whose industries had relatively low ratios and three whose industry ratios

\textsuperscript{132} Cohen, Nelson, and Walsh, "Protecting Their Intellectual Assets."

\textsuperscript{133} To calculate these, I simply averaged the reported effectiveness of patenting and secrecy for product and process innovations in each industry. While this method does implicitly assume that every industry invents the same number of products as it does processes, the specific weighting is not critical to the experimental design as the outcome of this analysis was simply to group industries into three categories by their relative importance of secrecy.
were in the middle of the industry listing for further analysis based on their diversity and the likelihood that each would contain a nontrivial number of patent classes. I also calculated the three additional metrics for each selected sector, for each of the high, average and low secrecy to patent ratio sector groupings, as well as for the sample as a whole to ensure that the sectors themselves fell into high, average and low secrecy to patent ratio classifications as intended. Once the nine sectors were selected, the next task was to identify patent classes associated with those sectors.

In order to associate patent classes with industry sectors, I used the 2005 concordance between the U.S. Patent Classification (USPC) system and the Office of Technology Assessment and Forecasting's (OTAF) product field sequence numbers based on the 1972 Standard Industrial Classification system (SIC). I began by surjectively mapping OTAF product field sequence numbers onto the selected sectors using the descriptive text accompanying the 2005 concordance. I then recorded each patent class in exclusive concordance with field codes associated with a single sector using a MATLAB program I created (see Appendix A2.1). Once this surjective composition was complete, time series data could be collected on patent applications in each of the remaining classes.

134 To estimate the overall effectiveness of patenting and secrecy in each sector, I used an average of the industry metrics weighted by the number of firms per industry in Cohen et al. While this calculation makes several assumptions including that each firm patents the same number of inventions and that the distribution of industries in industry sectors is identical between Cohen et al, Pennsylvania, and New Jersey, the specific weighting is not critical to the experimental design as the outcome of this analysis was simply to check that the sectors behaved similarly to the industries that composed them. To calculate the third metric, I took the ratio of the first to the second, giving the ratio of the percent of innovations in the sector for which secrecy is effective to the percent of innovations in the sector for which patenting is effective. While one might be tempted to calculate the third metric as the average of each industry’s secrecy to patent ratio weighted by the number of firms per industry, that calculation is inaccurate as it convolves the percent of innovations in the sector for which secrecy is effective with the industrial origin of those innovations.

135 United States Patent and Trademark Office, “Concordance Between the U.S. Patent Classification (USPC) System and the Standard Industrial Classification System (SIC),” Patent Technology Monitoring Division Research Publications, http://www.uspto.gov/products/catalog/ptmd/research.jsp#heading-3 (accessed on August 5, 2010). I used the 2005 concordance instead of a compilation of previously published concordances, the 2005 is the most recent, as it was the only one available in a manipulable format.

136 This meant that all patent classes that had mapped to multiple sectors or had mapped to any field code that was not included in a selected sector were eliminated from future analysis. This elimination was necessary to ensure that time series data was not collected on patent classes whose application rates could be affected by sectors not included in the study. Since the concordance is at the national level, it is possible that some patent classes were excluded unnecessarily as the analysis focuses exclusively on Pennsylvania and New Jersey. However, this meant that the patent classes which remained are not only exclusively associated with a certain sector in New Jersey and Pennsylvania, but in the entire United States.
4.2 Data Acquisition and Difference in Difference Regressions

The first step in acquiring time series data on patent applications in the remaining classes and on trade secrecy litigation in Pennsylvania and New Jersey was to identify which of the remaining classes had enough applications from March 15, 2001 to March 14, 2008 to be studied independently. To accomplish this I used the advanced search function of the Patent Application Full Text and Image Database (AppFT) maintained by the United States Patent Office. Specifically, I searched for patent applications whose application date was included in the above range for each remaining patent class with assignments in both Pennsylvania and New Jersey. Time series data was collected on each patent class for which at least one state had four times the number of applications as time periods and no state had fewer than three times the number of applications as time periods. In addition, time series data was collected on the set of all patent applications in each state, the set of all patent applications in each sector and the sets of all patent applications in high, average and low performing sectors. Observations were taken annually using years beginning on March 15 from 2001 to 2008. The title of each patent class was also recorded. Once time series data was collected for patent applications, the next task was to collect time series data for trade secrecy litigation.


138 An example search string for patent class *** in New Jersey would be "AS/NJ AND CCL/***/$ AND APD/3/15/2001->3/14/2008".

139 This lower bound on the number of applications per patent class in the two states eliminated patent classes for which there was not enough data to perform regressions.

140 An example search string identifying all patent applications assigned in a particular state over one year time period would be “AS/NJ AND APD/3/15/2001->3/14/2002”. Similarly, an search string identifying all patent applications assigned to a particular sector would have the form "AS/NJ AND APD/3/15/2001-3/14/2002 AND (CCL/****/$/ OR CCL/****2/$ OR ...)" where **** represents the ith patent class in the sector. The search string form for combining sectors is identical to that within a sector, except that the domain of patent classes expanded. It is also important to note that these values do not double count patents that are listed in more than one class, whereas a simple total of all patents in all classes would.

141 United States Patent and Trademark Office, “U.S. Classes by Number with Title Menu.”
I collected time series data for trade secrecy litigation in New Jersey and Pennsylvania state courts using the Lexis database, which contains all reported judicial opinions, and some unreported decisions, for the years it includes. Specifically, I searched the following courts for any case that contained the term "trade secre!" in any descriptive field from April 19, 1999 to April 18, 2009:

- Pennsylvania Supreme Court
- Pennsylvania Superior Courts
- Pennsylvania Commonwealth Courts
- Pennsylvania Courts of Common Pleas
- Pennsylvania Commonwealth Court of Judicial Discipline beginning in 1994
- New Jersey Supreme Court (includes former Court of Errors and Appeals)
- New Jersey Superior Court
- New Jersey Tax Court
- Unpublished New Jersey Superior Court Opinions beginning in 2005

Observations were collected annually using years beginning April 19 from 1999 to 2009 and grouped by state. This meant that all cases in a single state in a single year were taken as one observation. As this search would clearly identify cases that did not make trade secret claims because of its breath, I read each case identified and determine whether or not trade secret claims were made. This combination of breath of search and hand coding eliminated all type I and type II errors of the trade secret litigation search methodology not intrinsic to the Lexis database. Once all of the time series data was collected, the next task was to perform the regressions designed in Chapter 3.

As discussed in Chapter 3, the difference in difference regressions performed were of the form

---

142 I restricted my analysis to the state courts in Pennsylvania and New Jersey, specifically excluding the federal district courts in those states, because the federal districts can hear cases originating from any state and are not required, or necessarily even inclined, to apply their resident state’s trade secret law. Therefore the federal district courts in Pennsylvania were no more affected by the PAUTSA than was the United States District Court for the District of New Jersey.

143 The search string "trade secre!" includes all two-word phrases beginning "trade secre" such as trade secret, trade secrets, and trade secrecy.
\[ y_i = \beta_0 + (s_i)\beta_2 + (s_i)(f_i)\beta_3 + (t)\beta_4 + (t)(s_i)(f_i)\beta_5 + \varepsilon_i \]

where \( y_i \) is the \( i \)th observation, \( s_i \) is an indicator variable for the treated state, \( f_i \) is an indicator variable for the treatment, \( t \) is the time in years since the first period, \( \beta_j \) is the \( j \)th regression coefficient and \( \varepsilon_i \) is the error term.\(^{144}\) Recall that \( \beta_1 \) represents the baseline level of observation in the control state, \( \beta_2 \) represents the difference between the baseline level of observation in the treatment state and the control state, \( \beta_3 \) represents the initial discontinuity resulting from the treatment, \( \beta_4 \) represents the time trend in the observable variable and \( \beta_5 \) represents the difference in the time trend between treated observations and untreated observations. Regressions for which the observations were patent applications had 14 observations with treatment taking effect in the year beginning March 15, 2003 and regressions for which the observations were trade secret cases had 20 observations with treatment taking effect in the year beginning April 19, 2004.\(^{145}\)

Regressions were performed on all patent applications in each patent class for which time series data was collected, on all patent applications in each sector, on all patent applications in the sets of high, average and low performing sectors and on all patent applications assigned in Pennsylvania or New Jersey using a MATLAB\textsuperscript{®} program I created (see Appendix A2.2). A single regression was also performed on the number of trade secret cases in Pennsylvania or New Jersey. I recorded the regression coefficients, their t-statistics, the p-value's of their t-statistics, the Durbin-Watson (DW) statistic (\( d \)) and the p-value of \( d \) for each regression.

After collecting the regression data, I performed DW tests for positive and for negative first-order autocorrelation of the residuals at significance levels of 5% and 1%. In

\(^{144}\) In the regression given in equation 4.1, the observations are either the number of patent applications in a specific class or set of classes in a given year and state or the number of trade secret cases in a given year and state. The treatment in this regression is the knowledge of the potential existence of the Pennsylvania Uniform Trade Secret Act (PAUTSA) for patent applications and the existence of PAUTSA as law for trade secrecy litigation.

\(^{145}\) The reason for the varying number of observations was that there were seven time periods observed for patent applications and 10 time periods observed for trade secret cases both of which were observed in two states. See Subchapter 3.3 for a discussion of the choice of time periods.
the test for positive first-order autocorrelation, $d$ is compared to tabulated upper and lower bounds ($d_L$ and $d_U$) as follows.\footnote{\textit{d}_L and \textit{d}_U are functions of the number of observations, the number of regression coefficients, and the significance level of the test.}

- If $d < d_L$, the positive first-order autocorrelation of the residuals is statistically significant.
- If $d > d_U$, the lack of positive first-order autocorrelation in the residuals is statistically significant.
- If $d_L < d < d_U$, the test is inconclusive.

Similarly, in the test for negative first-order autocorrelation, $4-d$ is compared to the same bounds as follows:

- If $4-d < d_L$, the negative first-order autocorrelation of the residuals is statistically significant.
- If $4-d > d_U$, the lack of negative first-order autocorrelation in the residuals is statistically significant.
- If $d_L < 4-d < d_U$, the test is inconclusive.

In all tests performed, I used the tables from Savin and White to determine the appropriate upper and lower bounds.\footnote{N Eugene Savin and Kenneth J White, "The Durbin-Watson Test for Serial Correlation with Extreme Sample Sizes or Many Regressors," \textit{Econometrica} 45, no. 8, Econometrica (1977): 1989-96.}
Chapter 5: Results

Chapter 4 described the empirical method used in this study in replicable detail. Chapter 5 reports the results of the application of that method. It begins by describing the intermediate results of selecting sectors for further study and refining those sectors into patent classes. The chapter then presents the results from the difference in difference regressions on the trade secret litigation time series from New Jersey and Pennsylvania. Chapter 5 concludes with a brief presentation of the results from the Durbin-Watson tests for first-order autocorrelation.

5.1 Intermediate Results and Summary Statistics

As described in Subchapter 4.1, several intermediate steps were performed before the principal regressions. The first of these steps was an estimation of the effectiveness of secrecy and the effectiveness of patenting across product and process innovations, as well as the calculation of the ratio of those metrics, for each industry in Cohen et al.148 These estimates are found in Table A1.1 with summary statistics in Table 5.1. The mean effectiveness of secrecy ranges from a high of 73.47% of innovations in the miscellaneous chemicals industry to a low of 26.48% of innovations in the printing and publishing industry. Interestingly, the mean effectiveness of patenting ranges from a high of 44.36% of innovations in the medical equipment industry to a low of 18.27% of innovations in the electronic components industry. In fact, the mean effectiveness of secrecy is higher than the effectiveness of patenting in each industry. Specifically, the ratios of the mean effectiveness of secrecy to the mean effectiveness of patenting range from a low of 1.12 in the special purpose machinery industry to a high of 3.30 in the food industry. This implies that trade secrecy is uniformly more effective at protecting innovations than patents, even for industries traditionally thought of as relying very heavily on the patent system, such as pharmaceuticals. The averages of these three metrics for the entire set of industries

148 Cohen, Nelson, and Walsh, “Protecting Their Intellectual Assets.”
presented in Cohen et al. are 50.36%, 27.82%, and 1.90 respectively. From these results, I clustered industries likely to patent in the same classes, that also had similar ratios, into industry sectors.

### Table 5.1 Summary Statistics for Mean and Relative Effectiveness of Intellectual Property Protection Mechanisms Across Product and Process Innovations

<table>
<thead>
<tr>
<th>Intellectual Property Protection Mechanism</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trade Secret Industry</td>
<td>34</td>
<td>50.36%</td>
<td>9.48%</td>
<td>26.48%</td>
<td>73.47%</td>
</tr>
<tr>
<td>Trade Secret Industry Sector</td>
<td>9</td>
<td>51.04%</td>
<td>8.96%</td>
<td>35.34%</td>
<td>60.85%</td>
</tr>
<tr>
<td>Trade Secret Composition of Industry Sectors</td>
<td>3</td>
<td>52.83%</td>
<td>3.84%</td>
<td>50.14%</td>
<td>57.22%</td>
</tr>
<tr>
<td>Patent Industry</td>
<td>34</td>
<td>27.82%</td>
<td>7.30%</td>
<td>10.36%</td>
<td>44.36%</td>
</tr>
<tr>
<td>Patent Industry Sector</td>
<td>9</td>
<td>28.94%</td>
<td>7.06%</td>
<td>20.62%</td>
<td>43.18%</td>
</tr>
<tr>
<td>Patent Composition of Industry Sectors</td>
<td>3</td>
<td>29.79%</td>
<td>7.05%</td>
<td>22.71%</td>
<td>36.81%</td>
</tr>
<tr>
<td>Trade Secret to Patent Effectiveness Ratio</td>
<td></td>
<td>1.90</td>
<td>0.49</td>
<td>1.12</td>
<td>3.30</td>
</tr>
<tr>
<td>Trade Secret to Patent Effectiveness Ratio</td>
<td></td>
<td>1.82</td>
<td>0.36</td>
<td>1.32</td>
<td>2.24</td>
</tr>
<tr>
<td>Trade Secret to Patent Effectiveness Ratio</td>
<td></td>
<td>1.84</td>
<td>0.42</td>
<td>1.39</td>
<td>2.21</td>
</tr>
</tbody>
</table>

Notes: Column I is organized by intellectual property mechanism across all industries in Cohen, et al., the industry sectors chosen for further analysis, and the compositions of those sectors clustered by their respective trade secret patent effectiveness ratios. The trade secret to patent effectiveness ratios in Column III were calculated using weights based on the number of firms in each industry and are therefore not equal to the ratio of the trade secret and patent values in Column III.

I created nine industry sectors by clustering industries with similar trade secret to patent effectiveness ratios that were also likely to patent in the same classes. I further classified the sectors as either having a high, average, or low importance of trade secrecy based on the trade secret to patent effectiveness ratios of the industries that compose them. See Figure 5.1 for a visual representation of this process. Specifically, I grouped the semiconductors and related equipment, electronic components and communication equipment industries into the electronics sector, the basic chemicals, miscellaneous chemicals and plastic resins industries into the chemical sector, and the mineral products, concrete, cement and lime, and glass industries into the inorganic materials sector. Industries that were taken as sectors in and of themselves included the rubber and plastic
industry, the precision instruments industry, the petroleum industry, the metal products industry, the drugs (pharmaceuticals) industry, and the electrical equipment industry. With respect to the importance of trade secrecy, the electronics, rubber and plastic, and precision instruments sectors were classified as having a high relative importance of trade secrecy, the chemicals, inorganic materials, and petroleum sectors were classified as having an average relative importance of trade secrecy and the metal products, pharmaceuticals and electrical equipment sectors were classified as having a low relative importance of trade secrecy.

**Figure 5.1 Classification of Industries into Industry Sectors**

<table>
<thead>
<tr>
<th>Industry</th>
<th>Industry Sector</th>
<th>Relative Importance of Trade Secrecy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semiconductors and Related Equipment</td>
<td>Electronic Components</td>
<td>Electronics (non-Computer)</td>
</tr>
<tr>
<td></td>
<td>Communication Equipment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rubber and Plastic</td>
<td>Rubber and Plastic</td>
</tr>
<tr>
<td></td>
<td>Precision Instruments</td>
<td>Precision Instruments</td>
</tr>
<tr>
<td>Basic Chemicals</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemicals, nec</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miscellaneous Chemicals</td>
<td>Chemicals (non-Petroleum)</td>
<td></td>
</tr>
<tr>
<td>Plastic Resins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mineral Products</td>
<td>Inorganic Materials (non-Metal)</td>
<td>Average</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Petroleum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Metal Products</td>
<td>Metal Products (non-Steel)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drugs</td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electrical Equipment</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** Column I is a list of the industries from Cohen et al. chosen for further analysis. Column II is a list of the sectors I clustered these industries into for further analysis. Column III is a qualitative representation of the relative importance of trade secrecy for the chosen industries and industry sectors.
The effectiveness of secrecy and of patenting across product and process innovations, as well as the ratio of these metrics, were calculated for each sector and composition of sectors introduced above. The results can be found in Table A1.2 with summary statistics in Table 5.1. The mean effectiveness of secrecy ranges from a high of 60.85% of innovations in the pharmaceuticals sector to a low of 35.34% of innovations in the electrical equipment sector. For patenting innovations, the mean effectiveness ranges from a high of 43.18% of innovations in the pharmaceuticals sector to a low of 20.62% of innovations in the electronics sector. In addition, the ratios of the mean effectiveness of secrecy to the mean effectiveness of patenting range from a low of 1.32 for the electrical equipment sector to a high of 2.24 for the electronics sector. More importantly, the sectors classified as having a high relative importance of trade secrecy have higher trade secret to patent effectiveness ratios than the sectors classified as having an average relative importance of trade secrecy. Likewise, the sectors classified as having an average relative importance of trade secrecy have higher trade secret to patent effectiveness ratios than the sectors classified as having a low relative importance of trade secrecy. However, the trade secret to patent ratios across the three compositions are not well distributed. Specifically, the ratios are 2.21, 1.92 and 1.39 for the compositions of sectors with high, average and low trade secret to patent ratios. However, as one moves from industry to industry sector and then eventually to compositions of those sectors, the mean values for trade secret effectiveness, patent effectiveness and their weighted ratio change significantly less than their respective standard deviations. This indicates at the industry sectors, when taken together, represent an unbiased sampling with respect effectiveness of intellectual property protection. It is also worth mentioning that the number of firms contained in Cohen et al.’s sample for each sector is between 20 and 70, with two notable exceptions. Specifically, the sample size of firms in the chemical sector is 176, whereas the sample size of firms in the metal products sector is only eight. This imbalance is reflected in the sector compositions as the number of firms in each are 101, 265 and 113 in descending importance of trade secrecy. After I identified industrial sectors with diverse intellectual property protection characteristics, I surjectively mapped patent classes onto those sectors.
By surjectively mapping patent classes onto industry sectors, one can study the patenting behavior of those sectors by examining time series data on the associated patent classes. As a first step, I matched industry sectors with Office of Technology Assessment and Forecasting (OTAF) sequence numbers so that I could then use the 2005 concordance between the U.S. Patent Classification (USPC) system and the OTAF sequence numbers to surjectively map patent classes onto industry sectors. See Figure A1.1 for a visual representation of the OTAF to industry sector mapping and Table A1.3 for a list of patent classes exclusively associated with the aforementioned industry sectors. Table A1.3 also contains the total number of patent applications in Pennsylvania and New Jersey for each associated patent class, industry sector, and composition of sectors previously discussed from March 15, 2001 to March 14, 2008 and Table 5.2 provides summary statistics for this data. The number of patent classes per sector and the number of patent applications per class, sector, and composition of sectors vary widely. Specifically, the number of patent classes per sector ranges from a low of zero for the pharmaceuticals sector to a high of 25 for the chemicals sector. In addition, the number of patent applications per class ranges from a high of 422 for multiplex communications in New Jersey to a low of zero in 27 out of 156 observations. This phenomenon is also observed at the sector level, where the number of applications ranges from a high of 1204 for the electronics sector in New Jersey to a low of zero for the rubber and plastics sector in Pennsylvania. As reflected above, New Jersey has nearly twice as many patent applications over the observation period as Pennsylvania. Specifically, there are 10,461 patent applications in New Jersey over the observation period, but only 5,899 applications in Pennsylvania. This effect is even more pronounced in the sample of patent classes associated with the selected industry sectors; New Jersey has 2,802 applications whereas Pennsylvania only has 1,281. Another interesting trend is that there were approximately eight times as many patent applications in sectors associated with a high relative importance of trade secrecy than in sectors associated with a low relative importance. Using the elimination criteria outlined in Subchapter 4.2, I collected time series data on all remaining patent classes at the class, sector, and composition of sectors levels.

149 The pharmaceuticals sector is notably missing from table A1.3 because it was not uniquely associated with one or more patent classes.
Table 5.2 Summary Statistics for Total Number of Patent Applications per State for Mapped Patent Classes in Pennsylvania and New Jersey for Seven Years Beginning March 15, 2001

<table>
<thead>
<tr>
<th>Level of Analysis</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patent Class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>156</td>
<td>29.44</td>
<td>61.00</td>
<td>0</td>
<td>422</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>78</td>
<td>19.64</td>
<td>35.93</td>
<td>0</td>
<td>140</td>
</tr>
<tr>
<td>New Jersey</td>
<td>78</td>
<td>39.23</td>
<td>77.50</td>
<td>0</td>
<td>422</td>
</tr>
<tr>
<td>Industry Sector</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>255.94</td>
<td>366.98</td>
<td>0</td>
<td>1204</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>8</td>
<td>161.13</td>
<td>245.46</td>
<td>0</td>
<td>730</td>
</tr>
<tr>
<td>New Jersey</td>
<td>8</td>
<td>350.75</td>
<td>455.83</td>
<td>3</td>
<td>1204</td>
</tr>
<tr>
<td>Composition of Industry Sectors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>680.83</td>
<td>646.01</td>
<td>118</td>
<td>1775</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>3</td>
<td>427.67</td>
<td>464.69</td>
<td>118</td>
<td>962</td>
</tr>
<tr>
<td>New Jersey</td>
<td>3</td>
<td>934.00</td>
<td>796.94</td>
<td>190</td>
<td>1775</td>
</tr>
</tbody>
</table>

Notes: Column I is organized by level of analysis across combined and separate geographies. The values for the "Total" rows in Column II are twice as large as for the individual geographies as there is an observation for each state.

Collecting time series data on patent applications in each class, industry sector, and composition of industry sectors that met the elimination criteria was the final step in preparation for performing difference in difference regressions to assess the effect of the PAUTSA on knowledge disclosure in Pennsylvania.\textsuperscript{150} This time series data can be found in Table A1.4, with summary statistics in Table 5.3, and was acquired using the procedure outlined in Subchapter 4.2. Only five of the original nine industry sectors are represented by enough patent applications over the observation period to be included in the regressions, and only four of the remaining sectors contained individual classes with enough applications. Specifically, the electronics, precision instruments, chemicals and electrical equipment sectors are represented by enough patent applications to support the study of individual associated patent classes. They contained seven, three, two and one

\textsuperscript{150} Trade secret litigation time series data must also be collected.
uniquely associated patent class respectively. The metal products sector is also included, but none of its associated patent classes contained enough patent applications to be included in the regressions themselves. Fortunately, there is at least one representative sector and individual patent class categorized as having a relatively high importance of trade secrecy, average importance of trade secrecy and low importance of trade secrecy. In addition, the trade secret litigation time series data can be found in Table A1.5. The number of cases is fewer than expected, especially in New Jersey. In fact, the time series would not have passed the elimination criteria used to select patent application data for regression due to the low number of cases in New Jersey. However, there does appear to be a negative correlation with time in the Pennsylvania time series. The above time series data enabled me to perform difference in difference regressions to assess the effect of the PAUTSA on knowledge disclosure in Pennsylvania.

Table 5.3 
Summary Statistics for Total Number of Patent Applications per State for Non-Eliminated Patent Classes in Pennsylvania and New Jersey for Seven Years Beginning March 15, 2001

<table>
<thead>
<tr>
<th>Level of Analysis</th>
<th>N</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patent Class</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>26</td>
<td>128.92</td>
<td>93.33</td>
<td>26</td>
<td>422</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>13</td>
<td>83.00</td>
<td>43.55</td>
<td>26</td>
<td>140</td>
</tr>
<tr>
<td>New Jersey</td>
<td>13</td>
<td>174.85</td>
<td>108.07</td>
<td>71</td>
<td>422</td>
</tr>
<tr>
<td>Industry Sector</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>403.10</td>
<td>400.23</td>
<td>43</td>
<td>1204</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>5</td>
<td>253.40</td>
<td>277.52</td>
<td>43</td>
<td>730</td>
</tr>
<tr>
<td>New Jersey</td>
<td>5</td>
<td>552.80</td>
<td>476.84</td>
<td>48</td>
<td>1204</td>
</tr>
<tr>
<td>Composition of Industry Sectors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>6</td>
<td>680.83</td>
<td>646.01</td>
<td>118</td>
<td>1775</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>3</td>
<td>427.67</td>
<td>464.69</td>
<td>118</td>
<td>962</td>
</tr>
<tr>
<td>New Jersey</td>
<td>3</td>
<td>934.00</td>
<td>796.94</td>
<td>190</td>
<td>1775</td>
</tr>
</tbody>
</table>

Notes: Column I is organized by level of analysis across combined and separate geographies. The values for the "Total" rows in Column II are twice as large as for the individual geographies as there is an observation for each state.

Although the chemicals sector only contains two associated patent classes that were not eliminated, none of the other patent classes or sectors associated with an average relative importance of trade secrecy meet the criteria for time series data collection. Beyond this, the chemicals sector unexpectedly dominates the time series of all sectors for which trade secrecy is of relative importance because of the unexpectedly low patent application levels in patent classes exclusively associated with the inorganic materials and petroleum sectors.
5.2 Difference in Difference Regressions

Difference in difference regressions were performed using the model developed in Subchapter 3.4 and the time series data described in Subchapter 5.1. The resulting regression coefficients can be found in Tables 5.4 and 5.5 for patent application time series and Table 5.6 for the trade secret litigation time series. While the regression on trade secret litigation yielded no significant results, the regressions on patent application data proved more informative. Specifically, there were treatment coefficients indicating significant substitution effects between patent applications and trade secrets for at least one regression at each of the patent class, industry sector and composition of industry sector levels including at the state level. The next several paragraphs present the aggregate behavior of the best-fit coefficients for the time series data on patent applications. The subchapter concludes by briefly describing the best-fit coefficients for the trade secret litigation time series.

In the majority of cases, the best-fit coefficients resulting from difference in difference regressions of the patent application time series data displayed the expected sign and magnitude. However significance, even at the 5% level, is only seen in approximately one third of the coefficients. The sign and relative magnitude of the difference in difference estimates of the effect of the treatment on the treated can be found in Table 5.7. As in Subchapter 3.5, the presentation of the best-fit coefficients begins with the coefficients associated with controlling factors and concludes with the treatment coefficients.

$\beta_1$, the best-fit coefficient corresponding to the baseline level of observation in the control state, is positive in all but one regression and is not as close to the baseline level of observation in the control state as expected.\(^{152}\) As a result, $\beta_1$ only shows significance at the 5% level or better in seven out of 23 regressions. $\beta_2$, the best-fit coefficient corresponding

\(^{152}\) The regression is on time series data for the patent class "Data processing: measuring, calibrating, or testing" in the precision instruments sector. This regression has several unusual features that will be discussed in the following paragraphs.
to the difference between the baseline level of observation in the treatment state and the control state, is of the correct\textsuperscript{153} sign in all but two regressions, one of which is the regression with a negative $\beta_1$, but did not generally result in an accurate baseline observation in the treatment state.\textsuperscript{154} As one would expect from this, $\beta_2$ only shows significance at the 5\% level or better in three out of 23 regressions. $\beta_4$, the best-fit coefficient corresponding to the yearly growth rate of the observable in the control state, is positive in all cases and on the order of, if not larger than, $\beta_1$ in the majority of cases as expected. Not surprisingly, $\beta_4$ is significant at the 1\% level in 21 out of 23 regressions and that the .1\% level in 18 of the regressions.

$\beta_3$, the best-fit coefficient corresponding to the difference in difference estimate of the immediate effect of the treatment on the treated, is negative as expected in all but two regressions, one of which is the regression with the negative $\beta_1$ and $\beta_2$ with the incorrect sign.\textsuperscript{155} In addition, $\beta_3$ is generally less than or on the order of 50\% of the model’s estimation of the untreated level at the time of treatment, as expected.\textsuperscript{156} However, there are six notable exceptions where $\beta_3$ is greater than or on the order of 75\% of the model’s estimation of the untreated level at the time of treatment.\textsuperscript{157} As one would expect with such exceptions, $\beta_3$ is significant at the 5\% level or better in only 5 out of the 23 regressions. Surprisingly, four of the significant coefficients are also four out of the five coefficients with exceptional magnitude.\textsuperscript{158} In addition, $\beta_3$ varies as expected with respect to the relative importance of trade secrecy, namely increasing in relative magnitude while remaining

\begin{footnotesize}
\begin{enumerate}
\item\textsuperscript{153} I use the term “correct” instead of expected here because it is clear from the data whether Pennsylvania has a larger or smaller original observation level than New Jersey.
\item\textsuperscript{154} The baseline observation in treatment state is $\beta_1 + \beta_2$. The additional regression which displays an incorrect sign for $\beta_2$ is on time series data for the total number of patent applications in all sectors that were associated with a high relative importance of trade secrecy.
\item\textsuperscript{155} The other regression with a positive $\beta_3$ was on time series data for the patent class “Telecommunications” in the electronics sector.
\item\textsuperscript{156} The untreated level at the time of treatment as calculated by the model is $\beta_1 + \beta_2 + 2\beta_4$.
\item\textsuperscript{157} These six regressions are based on time series data for the patent classes “Pulse or digital communications” and “Interactive video distribution systems” in the electronics sector, the “Surgery” patent class in the precision instruments sector, the electrical equipment and metal products sectors and the composition of sectors that are associated with a low relative importance of trade secrecy.
\item\textsuperscript{158} The $\beta_3$ coefficients that have large negative magnitudes but are not significant correspond to the regressions on time series data for the metal products sector and the patent class “Surgery” in the precision instruments sector. The $\beta_3$ that was significant but did not have a relatively large magnitude was for the regression on time series data for all patent applications in the sample.
\end{enumerate}
\end{footnotesize}
negative as one moves from regimes with high trade secret to patent effectiveness ratios to those with lower ratios.

$\beta_5$, the best-fit coefficient corresponding to the difference in difference estimate of the additional effect of the treatment on the treated in each time period, is negative as expected in all but four regressions and smaller in magnitude than $\beta_4$ in all regressions in which $\beta_5$ was negative, also as expected. Interestingly, the four regressions in which $\beta_5$ is positive are four out of the five in which $\beta_3$ has a relatively large magnitude. In fact, in the other regression in which $\beta_3$ has a relatively large magnitude, $\beta_5$ has a small negative value. Similarly, $\beta_5$ has unexpectedly small negative, or even positive, values in regimes of low trade secret importance, but behaves as expected in high and medium ranges by increasing in magnitude while remaining negative as trade secrecy importance decreases. As one might expect with this level of variance, $\beta_5$ is significant at the 5% level or better in only 4 out of 23 regressions. In addition to regressions on time series data of patent applications, I performed a regression on trade secret litigation time series data.

---

159 The regression that has a $\beta_3$ with a relatively a large negative magnitude but does not have a negative $\beta_5$ is based on time series data from the metal products sector.

160 The p-value of $\beta_5$ in the regression consisting of patent application data for all applications in the sample is 0.0516 and so is not included in this total. However, I believe it is worth mentioning due the generality of the specific time series and the margin by which the p-value is considered insignificant.
Table 5.4  Difference in Difference Regressions of Patent Application Time Series Data at the State and Composition of Industry Sector Levels

<table>
<thead>
<tr>
<th>Set of Patent Classes</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3$</th>
<th>$\beta_4$</th>
<th>$\beta_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Patent Applications</td>
<td>324.2180*</td>
<td>-193.2531</td>
<td>-368.7649</td>
<td>390.0702***</td>
<td>-68.2702</td>
</tr>
<tr>
<td></td>
<td>(100.9330)</td>
<td>(138.3844)</td>
<td>(226.4565)</td>
<td>(21.9172)</td>
<td>(54.7775)</td>
</tr>
<tr>
<td>All Patent Applications in Sample</td>
<td>92.3910*</td>
<td>-59.7068</td>
<td>-116.6842*</td>
<td>102.6316***</td>
<td>-25.9316</td>
</tr>
<tr>
<td></td>
<td>(21.5819)</td>
<td>(-29.5899)</td>
<td>(48.4219)</td>
<td>(5.9646)</td>
<td>(11.7127)</td>
</tr>
<tr>
<td></td>
<td>(18.4965)</td>
<td>(25.3597)</td>
<td>(41.4994)</td>
<td>(5.1160)</td>
<td>(10.0383)</td>
</tr>
<tr>
<td>Identified Sectors with Average Secrecy to Patent Ratio</td>
<td>51.8346***</td>
<td>-52.6241***</td>
<td>-22.0105</td>
<td>22.5789***</td>
<td>-7.7789</td>
</tr>
<tr>
<td></td>
<td>(8.2558)</td>
<td>(11.3191)</td>
<td>(18.5229)</td>
<td>(2.2835)</td>
<td>(4.4805)</td>
</tr>
<tr>
<td>Identified Sectors with Low Secrecy to Patent Ratio</td>
<td>0.9850</td>
<td>-0.8446</td>
<td>-14.5404**</td>
<td>8.7193***</td>
<td>0.3807</td>
</tr>
<tr>
<td></td>
<td>(1.7676)</td>
<td>(2.4235)</td>
<td>(3.9659)</td>
<td>(0.4889)</td>
<td>(0.9593)</td>
</tr>
</tbody>
</table>

* $p < .05$; ** $p < .01$; *** $p < .001$. 
Table 5.5  Difference in Difference Regressions of Patent Application Time Series Data at the Industry Sector and Patent Class Levels

<table>
<thead>
<tr>
<th>Patent Class or Set of Classes</th>
<th>$\beta_1$</th>
<th>$\beta_2$</th>
<th>$\beta_3$</th>
<th>$\beta_4$</th>
<th>$\beta_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Electronics (non-Computer)</strong></td>
<td>34.0000</td>
<td>5.0000</td>
<td>-61.4000</td>
<td>46.0000***</td>
<td>-10.2000</td>
</tr>
<tr>
<td></td>
<td>(16.3854)</td>
<td>(22.4652)</td>
<td>(36.7628)</td>
<td>(4.5321)</td>
<td>(8.8925)</td>
</tr>
<tr>
<td>Active solid-state devices (e.g., transistors, solid-state diodes)</td>
<td>6.1654</td>
<td>4.9574</td>
<td>-11.5228</td>
<td>5.7544**</td>
<td>-0.0544</td>
</tr>
<tr>
<td></td>
<td>(5.3720)</td>
<td>(7.3653)</td>
<td>(12.0529)</td>
<td>(1.4859)</td>
<td>(2.9155)</td>
</tr>
<tr>
<td></td>
<td>(5.3589)</td>
<td>(7.3473)</td>
<td>(12.0233)</td>
<td>(1.4822)</td>
<td>(2.9083)</td>
</tr>
<tr>
<td>Pulse or digital communications</td>
<td>8.9098*</td>
<td>3.9323</td>
<td>-19.2421*</td>
<td>4.3158***</td>
<td>1.6842</td>
</tr>
<tr>
<td></td>
<td>(3.0052)</td>
<td>(4.1202)</td>
<td>(6.7425)</td>
<td>(0.8312)</td>
<td>(1.6309)</td>
</tr>
<tr>
<td>Optical communications</td>
<td>-3.9098</td>
<td>4.0677</td>
<td>-0.3579</td>
<td>4.6842**</td>
<td>-3.5842</td>
</tr>
<tr>
<td></td>
<td>(4.6303)</td>
<td>(6.3484)</td>
<td>(10.3887)</td>
<td>(1.2807)</td>
<td>(2.5129)</td>
</tr>
<tr>
<td></td>
<td>(2.3133)</td>
<td>(3.1716)</td>
<td>(5.1901)</td>
<td>(0.6398)</td>
<td>(1.2554)</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>4.6316</td>
<td>-3.5263</td>
<td>7.0947</td>
<td>8.7895***</td>
<td>-5.0895</td>
</tr>
<tr>
<td></td>
<td>(4.3880)</td>
<td>(6.0162)</td>
<td>(9.9389)</td>
<td>(1.2253)</td>
<td>(2.4041)</td>
</tr>
<tr>
<td>Interactive video distribution systems</td>
<td>8.7293</td>
<td>1.1303</td>
<td>-29.0596*</td>
<td>1.2807</td>
<td>9.1193**</td>
</tr>
<tr>
<td></td>
<td>(1.9706)</td>
<td>(0.1861)</td>
<td>(-2.9238)</td>
<td>(1.0453)</td>
<td>(3.7932)</td>
</tr>
<tr>
<td><strong>Precision Instruments</strong></td>
<td>5.6015</td>
<td>-11.2155</td>
<td>-19.5860</td>
<td>25.2281***</td>
<td>-7.9281*</td>
</tr>
<tr>
<td></td>
<td>(5.9279)</td>
<td>(8.1274)</td>
<td>(13.3000)</td>
<td>(1.6396)</td>
<td>(3.2171)</td>
</tr>
<tr>
<td></td>
<td>(4.6352)</td>
<td>(6.3798)</td>
<td>(10.4400)</td>
<td>(1.2870)</td>
<td>(2.5253)</td>
</tr>
<tr>
<td>Prosthesis (i.e., artificial body members), parts thereof, or aids ...</td>
<td>1.7519</td>
<td>-5.1028</td>
<td>-2.0491</td>
<td>7.7018***</td>
<td>-4.5018**</td>
</tr>
<tr>
<td></td>
<td>(2.4274)</td>
<td>(3.3281)</td>
<td>(5.4462)</td>
<td>(0.6714)</td>
<td>(1.3174)</td>
</tr>
<tr>
<td>Data processing: measuring, calibrating, or testing</td>
<td>-0.736</td>
<td>0.6140</td>
<td>0.5228</td>
<td>5.2456***</td>
<td>-2.8456</td>
</tr>
<tr>
<td></td>
<td>(2.6003)</td>
<td>(3.5651)</td>
<td>(5.8341)</td>
<td>(0.7192)</td>
<td>(1.4112)</td>
</tr>
<tr>
<td><strong>Chemicals (non-Petroleum)</strong></td>
<td>52.6391***</td>
<td>-54.9373***</td>
<td>-20.5018</td>
<td>20.5965***</td>
<td>-6.4965</td>
</tr>
<tr>
<td></td>
<td>(7.8092)</td>
<td>(10.7068)</td>
<td>(17.5209)</td>
<td>(2.1600)</td>
<td>(4.2381)</td>
</tr>
<tr>
<td>Organic compounds -- part of the class 532-570 series (548)</td>
<td>29.1128***</td>
<td>-29.8321***</td>
<td>-5.0807</td>
<td>6.4386***</td>
<td>-2.0386</td>
</tr>
<tr>
<td></td>
<td>(4.4233)</td>
<td>(6.0645)</td>
<td>(9.9242)</td>
<td>(1.2234)</td>
<td>(2.4006)</td>
</tr>
<tr>
<td>Organic compounds -- part of the class 532-570 series (549)</td>
<td>8.4060**</td>
<td>-7.8622</td>
<td>-0.1439</td>
<td>1.9123</td>
<td>-0.7123</td>
</tr>
<tr>
<td></td>
<td>(3.3600)</td>
<td>(4.6068)</td>
<td>(7.5386)</td>
<td>(0.9294)</td>
<td>(1.8235)</td>
</tr>
<tr>
<td><strong>Metal Products (non-Steel)</strong></td>
<td>0.1729</td>
<td>0.2130</td>
<td>-3.5860</td>
<td>2.2281***</td>
<td>0.5719</td>
</tr>
<tr>
<td></td>
<td>(1.0461)</td>
<td>(1.4343)</td>
<td>(2.3471)</td>
<td>(0.2893)</td>
<td>(0.5677)</td>
</tr>
<tr>
<td><strong>Electrical Equipment</strong></td>
<td>0.8120</td>
<td>-1.0576</td>
<td>-10.9544*</td>
<td>6.4912***</td>
<td>-0.1912</td>
</tr>
<tr>
<td></td>
<td>(1.6535)</td>
<td>(2.2671)</td>
<td>(3.7099)</td>
<td>(0.4573)</td>
<td>(0.8974)</td>
</tr>
<tr>
<td>Electricity: measuring and testing</td>
<td>2.2481</td>
<td>-1.5639</td>
<td>0.5158</td>
<td>2.6316**</td>
<td>-1.5316</td>
</tr>
<tr>
<td></td>
<td>(2.1955)</td>
<td>(3.0102)</td>
<td>(4.9259)</td>
<td>(0.6073)</td>
<td>(1.1915)</td>
</tr>
</tbody>
</table>

Notes: Regressions consisting of multiple patent classes are bolded.
* $p < .05$; ** $p < .01$; *** $p < .001$.


The results of the difference in difference regression on the time series of trade secret litigation in Pennsylvania and New Jersey are almost all unexpected. $\beta_1$ is an exception and was very close to the control state's original observation. $\beta_2$ is the incorrect sign and its magnitude is far too small. $\beta_3$ is also unexpectedly negative. $\beta_4$ is very close to zero instead of being large and positive as expected. However, $\beta_5$ is positive as theoretically expected. Not surprisingly, none of the best-fit coefficients are significant at the 5% level of better.

<table>
<thead>
<tr>
<th>Table 5.6</th>
<th>Difference in Difference Regressions of Trade Secret Litigation Time Series Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_1$</td>
<td>$\beta_2$</td>
</tr>
<tr>
<td>0.8324</td>
<td>-0.1514</td>
</tr>
<tr>
<td>(.7573)</td>
<td>(0.8130)</td>
</tr>
</tbody>
</table>

* $p < .05$; ** $p < .01$; *** $p < .001$. 
Table 5.7  Sign and Relative Magnitude of the Treatment Coefficients from Difference in Difference Regressions of Patent Application Time Series Data

<table>
<thead>
<tr>
<th>Patent Class or Set of Classes</th>
<th>$\beta_5 / (\beta_1 + \beta_2 + 2\beta_3)$</th>
<th>$\beta_5 / \beta_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Patent Applications</td>
<td>-40.47%</td>
<td>-17.50%</td>
</tr>
<tr>
<td>All Patent Applications in Sample</td>
<td>-49.04%</td>
<td>-25.27%</td>
</tr>
<tr>
<td>Identified Sectors with High Secrecy to Patent Ratio</td>
<td>-45.53%</td>
<td>-25.98%</td>
</tr>
<tr>
<td>Identified Sectors with Average Secrecy to Patent Ratio</td>
<td>-49.61%</td>
<td>-34.45%</td>
</tr>
<tr>
<td>Identified Sectors with Low Secrecy to Patent Ratio</td>
<td>-82.71%</td>
<td>4.37%</td>
</tr>
<tr>
<td>Electronics (non-Computer)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active solid-state devices (e.g., transistors, solid-state diodes)</td>
<td>-50.91%</td>
<td>-0.95%</td>
</tr>
<tr>
<td>Multiplex communications</td>
<td>-29.74%</td>
<td>-61.41%</td>
</tr>
<tr>
<td>Pulse or digital communications</td>
<td>-89.61%</td>
<td>39.02%</td>
</tr>
<tr>
<td>Optical communications</td>
<td>-3.76%</td>
<td>-76.52%</td>
</tr>
<tr>
<td>Semiconductor device manufacturing: process</td>
<td>-6.20%</td>
<td>-54.03%</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>37.97%</td>
<td>-57.90%</td>
</tr>
<tr>
<td>Interactive video distribution systems</td>
<td>-233.96%</td>
<td>712.06%</td>
</tr>
<tr>
<td>Precision Instruments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td>-43.68%</td>
<td>-31.43%</td>
</tr>
<tr>
<td>Prosthesis (i.e., artificial body members), parts thereof, or</td>
<td>-73.65%</td>
<td>-13.70%</td>
</tr>
<tr>
<td>Data processing: measuring, calibrating, or testing</td>
<td>-17.00%</td>
<td>-58.45%</td>
</tr>
<tr>
<td>Data processing: measuring, calibrating, or testing</td>
<td>5.04%</td>
<td>-54.25%</td>
</tr>
<tr>
<td>Chemicals (non-Energy)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organic compounds -- part of the class 532-570 series (548)</td>
<td>-52.71%</td>
<td>-31.54%</td>
</tr>
<tr>
<td>Organic compounds -- part of the class 532-570 series (549)</td>
<td>-41.79%</td>
<td>-31.66%</td>
</tr>
<tr>
<td>Metal Products (non-Steel)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electric Equipment</td>
<td>-74.06%</td>
<td>25.67%</td>
</tr>
<tr>
<td>Electricity: measuring and testing</td>
<td>-86.01%</td>
<td>-2.95%</td>
</tr>
</tbody>
</table>

Notes: All sets of patent classes, as represented by industry sectors or compositions of sectors, are bolded in Column I. Column II is the ratio of the difference in difference estimate of the immediate effect of the treatment on the treated to the model's estimation of the untreated level at the time of treatment. Column III is the ratio of the difference in difference estimate of the effect of the treatment on the treated overtime to the baseline growth rate of the observation.

5.3 Durbin-Watson Tests for Autocorrelation

As introduced and described in Subchapters 3.4 and 4.2, the Durbin-Watson statistic tests for first-order autocorrelation of the residuals. I performed this test on all regressions whose results were presented in Subchapter 5.2 in order to determine if the significance

161 In this table, this is the ratio of the immediate drop in patent applications as a result of the PAUTSA to the modeled level of patent applications that Pennsylvania would have had if it had not implemented the PAUTSA.

162 In this table, this represents the percent reduction in the growth rate of patent applications as result of the PAUTSA.
levels found were accurate, overestimated, or underestimated. The Durbin-Watson statistics as well as the results of the associated tests for positive and negative autocorrelation at the 5% and 1% significance levels can be found in Table 5.8 and Table 5.9 for regressions on patent application time series data and on secret litigation time series data respectively. None of the regressions on either patent applications or trade secret litigation time series data display positive first-order autocorrelation of the residuals. In fact, only four out of the 48 tests for positive first-order autocorrelation of the residuals are indeterminate. On the other hand, only seven of the 48 tests for negative first-order autocorrelation of the residuals are not indeterminate. Of these, five show no negative autocorrelation and two confirm the existence of negative autocorrelation in the residuals. These results are unexpected, as difference in difference studies generally yield positive autocorrelation of the residuals and therefore require a correction for overestimated significance levels. Since there is no evidence of positive autocorrelation in the residuals of the regressions presented in Subchapter 5.2, there is no need for a correction.

163 The tests for positive first-order autocorrelation of the residuals that lead to an indeterminate result are for the 5% and 1% significance levels on the regressions corresponding to the patent class “Data: processing: measuring, calibrating, or testing” in the electronics sector and the set of all patent applications in Pennsylvania and New Jersey over the observation period.

164 The five tests that showed no negative first-order autocorrelation of the residuals are at the 5% and 1% significance levels for the regressions corresponding to the patent class “Data: processing: measuring, calibrating, or testing” in the electronics sector and the set of all patent applications in Pennsylvania and New Jersey over the observation period as well as the test at the 1% significance level for the patent class “Pulse or digital communications” in the electronics sector. The two tests that confirm negative first-order autocorrelation of the residuals are the 5% significance level for the regression on patent application data for the entire chemical sector and for patent application data on sectors for which trade secrecy is of average relative importance.
Table 5.8  Durbin-Watson Tests for First-Order Autocorrelation of the Residuals from Difference in Difference Regressions of Patent Application Time Series Data

<table>
<thead>
<tr>
<th>Patent Class or Set of Classes</th>
<th>DW</th>
<th>5% Sig.</th>
<th>1% Sig.</th>
<th>5% Sig.</th>
<th>1% Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Electronics (non-Computer)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active solid-state devices (e.g., transistors, solid-state diodes)</td>
<td>2.5351</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td>Multiplex communications</td>
<td>2.2625</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td>Pulse or digital communications</td>
<td>2.1557</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>None</td>
</tr>
<tr>
<td>Optical communications</td>
<td>2.4033</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td>Semiconductor device manufacturing: process</td>
<td>3.1208</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>2.9811</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td>Interactive video distribution systems</td>
<td>2.6922</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td><strong>Precision Instruments</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td>2.5938</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td>Prosthesis (i.e., artificial body members), parts thereof, or aids and accessories therefor</td>
<td>3.2268</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td>Data processing: measuring, calibrating, or testing</td>
<td>3.1926</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td><strong>Chemicals (non-Petroleum)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organic compounds -- part of the class 532-570 series (548)</td>
<td>3.4126</td>
<td>None</td>
<td>None</td>
<td>Present</td>
<td>Indeterminate</td>
</tr>
<tr>
<td>Organic compounds -- part of the class 532-570 series (549)</td>
<td>2.8610</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td><strong>Metal Products (non-Steel)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Electrical Equipment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Electricity: measuring and testing</td>
<td>2.3545</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td><strong>All Patent Applications</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>All Patent Applications in Sample</strong></td>
<td>2.2727</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td><strong>Identified Sectors with High Secrecy to Patent Ratio</strong></td>
<td>2.2833</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
<tr>
<td><strong>Identified Sectors with Average Secrecy to Patent Ratio</strong></td>
<td>3.4937</td>
<td>None</td>
<td>None</td>
<td>Present</td>
<td>Indeterminate</td>
</tr>
<tr>
<td><strong>Identified Sectors with Low Secrecy to Patent Ratio</strong></td>
<td>2.3296</td>
<td>None</td>
<td>None</td>
<td>Indeterminate</td>
<td>Indeterminate</td>
</tr>
</tbody>
</table>

Notes: Regressions consisting of multiple patent classes are bolded.
Table 5.9  Durbin-Watson Tests for First-Order Autocorrelation of the Residuals from the Difference in Difference Regression of Trade Secret Litigation Time Series Data

<table>
<thead>
<tr>
<th>$DW$</th>
<th>Positive Autocorrelation</th>
<th>Negative Autocorrelation</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5054</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>5% Sig.</td>
<td>1% Sig.</td>
</tr>
<tr>
<td></td>
<td>5% Sig.</td>
<td>Indeterminate</td>
</tr>
<tr>
<td></td>
<td>1% Sig.</td>
<td>Indeterminate</td>
</tr>
</tbody>
</table>
Chapter 6: Discussion

Chapter 5 presented the results of the difference in difference regressions of patent application and trade secret litigation time series data in Pennsylvania and New Jersey. Chapter 6 discusses the implications of those results and places them in the greater context of knowledge disclosure. The chapter begins with an examination of the trade secret litigation regression. It then explores the potential influence of the author's design choices on the intermediate results. This is followed by a discussion the aggregate properties displayed by the control coefficients and the dynamics of the treatment coefficients. The chapter concludes by discussing the effect of possibility of negatively autocorrelated residuals and provides insight into the general autocorrelation phenomena observed in difference in difference analyses.

6.1 Trade Secret Litigation

A difference in difference regression of the form presented in Equation 3.4 was performed on the trade secret litigation time series data presented in Table A1.5. However, quality of this regression is suspect for numerous reasons. First, the level of observation is too small to have passed the elimination criteria set forth for patent application time series data. These criteria had nothing to do with the nature of patent applications, but simply existed to ensure that there were enough observations in each time period for an effect to be distinguishable from the inherent noise in time series data.\footnote{See Subchapter 4.2 for a description of the selection criteria.} In fact, the time series data for New Jersey appears to be noise upon examination, with the average number of trade secret litigation cases per observation period equal to 1.1. On the other hand, the trade secret litigation time series data from Pennsylvania has a sufficient observation level; and upon examination, there appears to be a distinct drop in trade secret cases with time after the enactment of the PAUTSA. However, without a viable control, this observation is
The inadequacy of the underlying time series data, specifically the lack of a viable control, explains the highly unexpected results of the trade secret litigation regression as well as the fact that none of the coefficients are statistically significant.

Without reliable trade secret litigation results, it is impossible to estimate the effect of the PAUTSA on the aggregate number of trade secrets in Pennsylvania. While this is one of the goals of the study, it is a secondary one of academic curiosity rather than practical value. The ability to achieve the primary goal of assessing how the PAUTSA affected firms' knowledge disclosure strategies in Pennsylvania remains unaffected. As discussed in Subchapter 3.2, the implementation of the PAUTSA can only affect the number of patent applications by substitution with trade secret creation because the statute has no direct impact on patent law. Therefore, results of the regressions on the patent application time series data are sufficient to assess the effect of the PAUTSA on the knowledge disclosure strategies of firms in Pennsylvania.

6.2 Influence of the Empirical Method

While Chapter 3 is devoted to developing and justifying the empirical design, some of the specific methods choices are not addressed. This subchapter addresses the influence of those choices on the data identification and collection process. This is particularly important as the empirical method described in Chapter 4 required significant preparatory calculations and manipulations that produced a large volume of intermediate results. These intermediate results in turn determined the composition of the time series data that served as the basis for the primary difference in difference analyses. Therefore understanding the inherent biases is crucial to understanding the limitations of the method as well as the primary results.

One of the first steps in the empirical method presented in Chapter 4 was creating a surjective composition mapping patent classes onto Office of Technology Assessment and

---

166 See Subchapter 4.3 for a discussion of the choice of New Jersey as the optimal control state for Pennsylvania.
Forecasting (OTAF) sequence numbers onto industry sectors. The reason for doing this was to uniquely associate patent classes with industry sectors, for which information about the relative effectiveness and importance of trade secrets to patents is known, to facilitate a straightforward comparative analysis of the effect of the Pennsylvania Uniform Trade Secrets Act (PAUTSA) on patent classes, industry sectors and compositions of sectors with varying relative levels of importance of trade secrecy to patenting. However, by eliminating all patent classes not uniquely associated with a given industry sector, a significant amount of potentially useful data is lost. For example, as noted in Subchapter 5.1, every patent class associated with the pharmaceuticals sector was also associated with at least one other sector; with the result that no time series data on the pharmaceuticals sector was collected. This phenomenon can be avoided while still maintaining the association between the relative importance of trade secrecy and patent class by relaxing the uniqueness requirement for the mapping of patent classes onto OTAF sequence numbers and transforming the aforementioned association from a discrete mapping to a continuous function by defining the relative importance of trade secrecy for a patent class as the weighted average of the trade secret to patent effectiveness ratios of all industry sectors with which its associated. One could then perform a regression to determine the effect of the PAUTSA on the number of patent applications per year in Pennsylvania as a function of trade secret patent effectiveness ratio. However, the effort to create this mapping as well as the amount of information one could mine from it would constitute a thesis project in and of itself. Therefore, the above solution is beyond the scope of this thesis and, as a result, the time series data is more heavily weighted toward unique industry sectors.

The step in the empirical method immediately following the surjective mapping of patent classes onto industry sectors is the elimination of patent classes, industry sectors, and compositions of industry sectors that do not meet the minimum total number of patent applications over the observation period.167 As previously mentioned in multiple subchapters, the purpose of this screening is to ensure that each set of time series data has a high enough observation level to be able to distinguish an effect from noise. However, this

---

167 See Subchapter 4.2 for a description of the selection criteria.
elimination has the potential to bias the sample of patent classes toward those that were less affected by the PAUTSA. The reason for this potential bias is that this process could selectively eliminate patent classes whose application level was initially high, but fell precipitously as result of the PAUTSA. While this elimination may select for time series data to be collected on individual patent classes that were relatively less affected by the PAUTSA, this has little to no impact on the findings of this study for two reasons. The first is that the observations of all patent classes eliminated by this process are included in their respective industry sector and industry sector composition regressions. The second is that individual patent class results are not aggregated. Therefore potential biases in the set of individual patent classes do not filter through to the aggregate results and are not material to the conclusions of this thesis.

At both the final elimination and time series data collection stages, as represented by Tables A1.3 and A1.4, it is clear that industry sectors with higher secrecy to patent effectiveness ratios are exclusively associated with more patent classes, have a higher average of patents per class and have a higher percent of classes with over 100 patent applications over the observable period. One potential explanation for this would be that even though these industry sectors had large trade secret to patent effectiveness ratios, the overall effectiveness of patenting in these sectors was also greater than the overall effectiveness of patenting in the sectors with lower trade secret to patent effectiveness ratios. However, this is not the case. Nonetheless, there are many possible explanations for this that have nothing to do with the experimental design of this study including, but not limited to, the history of the respective industry sectors and their relative size. This phenomenon, in combination with the elimination of non-unique patent classes through the use of a surjective mapping, is the reason that the composition of all sectors associated with a high relative importance of trade secrecy is dominated by the electronics sector, the composition of all sectors associated with an average relative importance of trade secrecy is dominated by the chemicals sector, and there are relatively few observations in the composition of all sectors associated with a low relative importance of trade secrecy. Note that the elimination process discussed in the previous paragraph has no effect on this phenomenon as the process has no effect on composition level data.
6.3 Summary of Central Findings

With an understanding of how choices in the empirical design of this study influenced the composition of the time series data collected, it is now possible to discuss this study’s primary results. Namely, the effect of the PAUTSA on patent application rates in Pennsylvania. This subsection provides a summary of the central findings, leaving the detailed analysis of the regression coefficients to Subchapters 6.4 and 6.5.

A significant substitution effect between patent applications and trade secrets is displayed in at least one regression at the patent class, industry sector, composition of industry sectors and statewide levels. While the time series data for all patent applications submitted during the observation period does not exhibit significant immediate or sustained decline in patent application rates, the filtered sample of all patents associated with a selected industry sector, either with or without enough patents to have been selected for individual time series analysis, displays a significant immediate substitution effect and an insignificant sustained effect with a p-value of 0.516. As the sampling method used to create this subset of patent applications, whose regression is labeled “All Patent Applications in Sample,” does not contain a bias toward effectiveness of either trade secrecy or patenting, it acts as a filter reducing noise in the overall state-level sample. Therefore it makes sense that if there were an effect, its significance would be more easily observed in the sample of patent applications associated with the selected industry sectors. However, the fact that the effect was only observed in the filtered sample is an indication that it is relatively weak.

Among the compositions of industry sectors associated by level of tracing the importance, only the composition associated with a lesser importance of trade secrecy shows a significant substitution effect between patent applications and trade secrets. Specifically, this composition shows a significant immediate effect unaccompanied by a sustained reduction in the growth rate of trade secret applications. While significant effects in all compositions were expected, the strength of the effect was expected to increase the
importance of trade secrecy decreased in the corresponding sample. The fact that only the composition associated with the lowest relative importance of trade secrecy display in effect is in agreement with this prediction, and further suggests a relatively weak substitution effect. While there were significant effects among the individual industry sectors and patent classes, an analysis of them at this broad level does not provide any additional insight beyond further supporting the conclusion that there is a relatively weak substitution effect between patent applications and trade secrets in Pennsylvania as a result of the implementation of the PAUTSA. To understand the nuances of this effect, the control and treatment coefficients are analyzed in detail in the subchapters below.

6.4 Aggregate Properties of the Control Coefficients

Of the 115 best-fit coefficients, five for each of the 23 regressions performed, only 40 were statistically significant at the 5% level of better. While the discussion will not be restricted to these significant coefficients, as the exact coefficient values are less important than their signs and relative magnitudes, special note will be taken of them throughout the discussion. As in Subchapters 3.5 and 5.2, the discussion of the best-fit coefficients begins with the coefficients associated with controlling factors and concludes with the treatment coefficients; in this case in the following subchapter.

The fact that the $\beta_1$ coefficients, which correspond to the number of new patent applications in New Jersey in the original time period, are almost exclusively positive but have a magnitude further from their correspondence than expected means very little on its own. However, while the $\beta_2$ coefficients, which correspond to the difference between the number of patent applications in Pennsylvania and New Jersey the original time period, are the correct sign in 21 of the 23 regressions, the fact that $\beta_1 + \beta_2$ is also inaccurate for many of the regressions suggests a nonlinear relationship between the number of new patent applications per year and time. This is further supported by the fact that only 10 out of the 46 calculated $\beta_1$ and $\beta_2$ coefficients are significant at the 5% level or better, whereas 21 out of 23 and 18 out of 23 $\beta_4$ coefficients, which correspond to the yearly growth rate in the
number of patent applications in New Jersey, are significant at the 1% and .1% levels or better. The reason these results imply nonlinearity in the time series data are described in the paragraph below.

When a linear function, such as the right hand side of Equation 3.4, attempts to fit data using an ordinary least squares (OLS) procedure, the objective function seeks to minimize the sum of the squared residuals. When this data is nonlinear, the objective function is often minimized by closely fitting the longest stretch of time series data with a relatively constant growth rate at the expense of accuracy at the initial values. According to Romer’s theory of endogenous growth, the number of patent applications grows exponentially as a function of the existing codified knowledge and human capital with increasing marginal productivity.\textsuperscript{168} Using an ordinary least squares regression to fit exponentially growing data results in the intercepts of the fit significantly underestimating the initial values in the data. This is the phenomenon being observed in the control variables. While one could potentially resolve this issue by using the model in Equation 3.3, doing so would eliminate the ability to examine and application data on the patent class level.\textsuperscript{169} Since understanding the basic dynamics of patent application rates at the patent class level in response to the PAUTSA is one of the primary goals of the study, rather than measuring precise changes at the industry sector level, some accuracy was implicitly exchanged for the ability to investigate at this granularity by selecting Equation 3.4 over Equation 3.3. Therefore, while not all significant at the 5% level or better, the signs and relative magnitudes of the treatment coefficients contain valuable information about the effect of the PAUTSA on knowledge disclosure in Pennsylvania.

\textsuperscript{168} While Romer specifies that the aggregate number of patents, not new patent applications, grow at an exponential rate, this implies that the number of new patent applications do as well because the derivative of an exponential remains exponential. See Romer, "Endogenous Technological Change."

\textsuperscript{169} See Subchapter 3.4 for discussion of the selection of the model used in the study.
6.5 Dynamics of the Treatment Coefficients

Now that the properties of the regression and control coefficients are sufficiently understood, the treatment coefficients can be examined. The fact that the $\beta_3$ coefficients, which correspond to the difference in difference estimate of the immediate effect of the PAUTSA on the annual rate of new patent applications in Pennsylvania, are almost exclusively negative; and are less than or approximately 50% of the model’s estimation of the untreated level at the time of treatment, indicates that there is an immediate substitution effect between patent applications and trade secrets as a result of the PAUTSA. Since this effect was expected, the associated reasoning can be found in Subchapter 3.5. Similarly the $\beta_5$ coefficients, which correspond to the difference in difference estimate of the additional effect of the PAUTSA on the annual rate of new patent applications in Pennsylvania in each time period, are mostly negative and are smaller than $\beta_4$ in all regressions in which they are negative. This indicates a reduction in the annual growth rate of patent applications as a result of a sustained substitution effect with trade secrets. Since this effect was also expected, the associated reasoning can be found in Subchapter 3.5 as well. However, the primary exceptions to the above characterizations about the $\beta_3$ and $\beta_5$ coefficients are represented in the same six regressions.170

These regressions display negative $\beta_3$ coefficients whose magnitudes are greater than or on the order of 75% of the model’s estimation of the untreated level at the time of treatment in conjunction with $\beta_5$ coefficients that are either on the order of or greater than zero. If these $\beta_3$ coefficients were not associated with the unusually non-negative $\beta_5$ coefficients, they would not have been selected for discussion, aside from their statistical significance, because they would simply have been viewed as examples industry sectors and patent classes that were particularly sensitive to the PAUTSA. However, this traditional explanation does not suffice because the $\beta_5$ coefficients imply that the PAUTSA either did

170 Supra 159. Upon closer examination, these six exceptional results represent only four independent and usable regressions. Specifically, the composition of sectors that are associated with a low relative importance of trade secrecy is exclusively composed of the metal products and electrical equipment sectors; while the “Interactive video distribution systems” patent class is one of two for which the tradeoff between accuracy and granularity was too great. Supra 156.
not impact, or positively impacted, the annual rate of patent applications in Pennsylvania. However, there is an explanation for this strange effect. It is possible that firms who patent in the corresponding patent classes and are in the associated industry sectors originally overestimated the degree to which the PAUTSA would enhance the value of trade secret protection, accounting for the negative $\beta_3$ coefficients with unusually large magnitudes, and then after seeing the implementation of the PAUTSA by the courts in the following year, realized their error and consequently patented a portion of the previously secret innovations as well as continued to patent new innovations at a level more akin to the one represented by the majority of regressions; accounting for the non-negativity of $\beta_3$. While this, or any other economic explanation is not currently verifiable with the data at hand, the existence of a non-zero number of plausible economic explanations can guide future study.

One can assess the effect of the PAUTSA on firms with varying levels of relative trade secret importance by examining the best-fit coefficients corresponding to compositions of sectors associated with these levels as well as with the sectors themselves. As expected, the $\beta_3$ coefficients increase in magnitude as one moves from regimes with a high trade to patent ratio to those with ratios closer to one. However, the $\beta_3$ coefficients associated with the chemicals sector and for the composition of sectors for which trade secrecy is of average relative importance are much closer to those for sectors for which trade secrecy is of high relative importance than low relative importance. There are two primary reasons for this. First, as described in Subchapter 6.2, the patent classes associated with the chemicals sector dominate the set of classes that compose composition of sectors for which trade secrecy is of average relative importance. This means that the regression coefficients, as well as other intermediate and final results, in the chemicals sector and in the composition in which it resides are very similar. Second, the trade secret to patent effectiveness ratios for this sector and composition of sectors are much closer to the ratio for the composition of sectors associated with a high relative importance of trade secrecy than to the ratio for the composition of sectors associated with a low relative importance of trade secrecy. Specifically, the ratios from high to low relative importance are 2.21, 1.92,

---

171 See Subchapter 3.3 for a discussion of the timing of the announcement and implementation of the PAUTSA.
and 1.39. This implies that the behavior of the two compositions not associated with a low relative importance of trade secrecy will be similar. In fact, this similarity is also apparent in the $\beta_5$ coefficients, but to a lesser degree. However, the $\beta_5$ coefficients do not simply increase in magnitude as one moves from regimes of where trade secrecy is relatively important to those in which it is not. Specifically, the composition of sectors in which trade secrecy is of low relative importance has a small, positive $\beta_5$ instead of a large, negative $\beta_5$ as expected. While the superficial reason for this is that both the metal products and electric equipment sectors, which form the composition, are both members of the six regression notable exception. However, the underlying reason for this would then likely be that the short-term compensatory patent application creation discussed in the previous paragraph overwhelmed lasting downward effect of the PAUTSA on the growth rate of the annual rate of patent application creation in the short-term. If this were the case, one would expect that $\beta_5$ would be a large negative value in 10 years.

### 6.6 Potential Negative Autocorrelation of the Residuals

As presented in Subchapter 5.3, Durbin-Watson tests for first-order autocorrelation of the residuals were performed for each regression to determine whether or not a correction was necessary to account for inflated significance levels. The fact that the results of these tests indicate little to no possible positive first-order autocorrelation of residuals, but do indicate potential negative first-order autocorrelation in almost all residuals, is surprising at first because the literature reports that the majority of difference in difference analyses are plagued by positive autocorrelation, not negative.\(^\text{172}\) However, upon closer examination of the study design, the reason for potential negative autocorrelation becomes clear.

In economics, many phenomena are understood to exhibit exponential growth with very low growth rates.\(^\text{173}\) Therefore empirical studies, such as difference in difference papers, are often confronted with data resembling slow growing exponential curves. Recall

\(^{172}\) Bertrand, Duflo, and Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?"

\(^{173}\) See the discussion of economic growth in Subchapter 1.1.
the discussion of fitting a linear function to exponential growth in Subchapter 6.4. Due to the convexity of exponential growth curves, this implies that the signs of the residuals are predetermined depending on what region of the curve they are associated with and how much of the curve is captured in the time series. The majority of difference in difference studies are performed over time scales which are large compared to the dynamics of their phenomena. This, combined with the low growth rates observed in natural phenomena, means that the data often captures the beginning of the noticeable convexity in the growth. A fitting of this data results in a few initial observations with large, negative residuals and many sequential observations with small positive residuals. This is the likely cause of the positive autocorrelation seen in the majority of difference in difference studies. However, if the timescale is shorter relative to the dynamics of the phenomena, as it is in this study, one signify reduces the number of positive residuals because the portion of the curve being observed is less convex. In addition, this reduces the magnitude of the initial negative residuals is the slope of the fit decreases. This explains the potential existence of negatively autocorrelated residuals in the results of this study.

While a correction could be implemented to correct for the possibility of negatively autocorrelated residuals, it would not be valuable to this study. While researchers should implement corrections for positive first-order autocorrelation even if the results of the Durbin-Watson tests are inconclusive, this is not the case for negative first-order autocorrelation. The reason for this difference is that in the first case, the researcher is being conservative by possibly underestimating the significance of his or her findings to avoid overestimating significance levels. In the second case, implementing the correction when statistically significant evidence is not present would itself be an overestimation because negatively autocorrelated residuals underestimate, not overestimate, significance levels. In addition, the implicit trade-off between accuracy and granularity discussed in Subchapter 6.2 precludes performing the correction to enhance the accuracy of the results.

174 For a discussion of the choice of data and timescale, please see Subchapters 3.2 and 3.3.  
175 Bertrand, Duflo, and Mullainathan, "How Much Should We Trust Differences-in-Differences Estimates?"
Chapter 7: Conclusion

This thesis began by introducing the concept of trade secrecy in the context of endogenous growth theory and codified knowledge. Chapter 1 then demonstrated that while trade secrets are an important part of our economy, little empirical evidence exists on their behavior. Chapter 2 provided a primer on trade secret law in the United States and identified the value of understanding the effects of statutory trade secret law. Chapter 3 introduced the idea of difference in difference analysis and designed a study to examine the effect of the Pennsylvania Uniform Trade Secrets Act (PAUTSA) on patent application rates and trade secrecy litigation in the state. Chapter 4 detailed a patent class to industry sector mapping for data coding and described the linear regressions and statistical tests used to analyze the patent application and trade secret litigation time series data. Chapter 5 presented the results of this method and stated whether or not they were expected. Chapter 6 discussed these results and established that the PAUTSA caused a substitution of patent applications for trade secrets that increased in strength as the effectiveness of trade secrecy as an alternative for a given patent class decreased. Chapter 7 suggests techniques that would enhance this study as well as directions for future work.

7.1 Enhancing the Study Beyond the Scope of this Thesis

This thesis devoted the previous six chapters to developing an understanding of the effect of implementing statutory trade secret law on knowledge disclosure. Specifically, it investigated whether the enactment of the PAUTSA created a substitution effect between patent applications and trade secrets. A weak substitution effect was found as a result of this legislation that varied across industries and industry sectors as a function of the relative importance of trade secrecy to patenting at the state, industry sector and patent class levels. Nonetheless, there are always ways to improve the quality of an empirical study that were beyond the scope of the original research. For example, the idea of enhancing this study by using a continuous mapping strategy to leverage all available patent application data and provide a more accurate association between patent classes
and the relative importance of trade secrecy was presented in subchapter 6.2. Subchapter 7.1 presents two additional ideas that, while beyond the scope of this thesis, would enhance the significance and validity of the study herein. Specifically, they are to control for the effect of legal forum shopping and to create a synthetic control state based on aggregate patent history.

While the difference in difference analysis discussed in Chapter 6 suggests a substitution effect between trade secrets and patent applications, this may not have been the only contributing factor to the observed change in patent applications rates. Legal forum shopping, the idea that innovators may move to or away from a state because it has relatively stronger or weaker protective institutions, may have influenced the results as well. Traditional ways to disentangle these effects include examining patents exclusively associated with large corporate laboratories that have difficulty relocating as well as coding the patent application data by firm size, i.e. recording whether each patent application was made by an individual, small firm, or large firm, and then performing regressions to measure the effect. However, both of these methods suffer from the problem of indirect measurement, i.e. one is making an assumption about the relationship between firm size and mobility and relying on that assumption to assess mobility. In fact, the first of these two methods also significantly restricts the available data, which may reduce the significance of the results. A more direct way to account for forum shopping would be to record the date of the oldest patent or patent application assigned to each assignee of an application in the sample in the state in which that patent application was assigned. One could then perform a regression with this information to determine the effect of legal form shopping. However, this could only demonstrate a greater substitution effect in this specific case because the overall observed effect was a decrease in the number of patent applications in response to an increase in general intellectual property protection. Since this change in intellectual property protection would encourage innovators to come to Pennsylvania, any effect from legal forum shopping would increase the number of patent applications. This implies that the substitution effect would have to compensate for the

176 For example, if in the sample of patent applications in State A there was an application assigned to XYZ Corporation, one would record the date of the earliest patent or patent application assigned to XYZ.
legal forum shopping before any negative aggregate effect could be observed. While controlling for legal forum shopping could more accurately assess the magnitude of the effect of the PAUTSA on patent application rates, it would do little to change the low significance levels observed in the data.

Creating a synthetic control state using historical patent data from a combination of Massachusetts, New York, New Jersey and Texas could improve the accuracy and significance of the results. While a synthetic control state could not be created using patent application data, there is more than enough historical patent data to create a synthetic control state. Even though the validity of this control would rely on the assumption that the distribution of patent applications and patents across patent class is identical within a given state, this would likely provide a more specific control for patent applications than selecting a state based on macroeconomic conditions and its prevalence in the literature. New York, New Jersey, Massachusetts and Texas would be chosen as the “donor” states because they are the only four that still rely exclusively on the common law for trade secret protection. In combination, the implementation of a continuous patent class to industry mapping, a control for legal forum shopping and a synthetic control state would improve the accuracy, validity and significance of this study.

7.2 Future Directions in Trade Secret Research

As is the case with all empirical trade secret studies to date, and most explorations into emerging fields of research, one is left with more questions than answers after its completion. While this thesis demonstrated a substitution effect between patent applications and trade secrets resulting from the enactment of a statutory trade secret law, the question of the legal effect still remains. Specifically, how the Pennsylvania courts initially interpreted the statute, and how this interpretation has changed over time. More generally, the insight into the effects of statutory trade secret law provided by this thesis contribute to answering one of the most pressing intellectual property policy questions in

177 See Subchapter 3.3 for a discussion of the creation of synthetic control state using patent application data.
178 See Subchapter 3.3 for discussion of the selection of New Jersey as a control state for Pennsylvania.
the United States; should there be a national trade secrets act that defines trade secret protection in the United States? While this thesis provides a basis for understanding what the effect of such an act might be, there is still significant debate over the idea that trade secrets should be protected, and if so, why this protection should exist. Until empirical research reveals the economic impact of trade secrecy, i.e. whether trade secrecy has a beneficial effect on the United States economy as measured by G.D.P. and other common metrics or not, the theoretical debate over its effects is likely to continue.
### Appendix 1: Supplemental Tables and Figures

#### A1.1 Tables

**Table A1.1**  
Mean and Relative Effectiveness of Intellectual Property Protection Mechanisms  
Across Product and Process Innovations by Industry

<table>
<thead>
<tr>
<th>Industry</th>
<th>((S_{\text{product}} + S_{\text{process}})/2)</th>
<th>((P_{\text{product}} + P_{\text{process}})/2)</th>
<th>((S_{\text{product}} + S_{\text{process}})/)</th>
<th>((P_{\text{product}} + P_{\text{process}})/)</th>
<th>(N_{\text{firms}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food</td>
<td>57.19</td>
<td>17.33</td>
<td>3.30</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td>Textiles</td>
<td>62.18</td>
<td>22.61</td>
<td>2.75</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>Printing/Publishing</td>
<td>26.48</td>
<td>10.36</td>
<td>2.56</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>Metal, nec</td>
<td>65.83</td>
<td>25.84</td>
<td>2.55</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Semiconductors and Related Equipment</td>
<td>58.75</td>
<td>25.00</td>
<td>2.35</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Search/Navigation Equipment</td>
<td>46.30</td>
<td>20.96</td>
<td>2.21</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td>Rubber and Plastic</td>
<td>58.00</td>
<td>26.29</td>
<td>2.21</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Electronic Components</td>
<td>40.29</td>
<td>18.27</td>
<td>2.21</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Miscellaneous Chemicals</td>
<td>73.47</td>
<td>33.49</td>
<td>2.19</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>Plastic Resins</td>
<td>59.45</td>
<td>27.13</td>
<td>2.19</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>Mineral Products</td>
<td>47.50</td>
<td>22.22</td>
<td>2.14</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Precision Instruments</td>
<td>45.42</td>
<td>21.32</td>
<td>2.13</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Steel</td>
<td>39.00</td>
<td>18.75</td>
<td>2.08</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Concrete, Cement, Lime</td>
<td>49.50</td>
<td>24.25</td>
<td>2.04</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Communication Equipment</td>
<td>41.26</td>
<td>20.22</td>
<td>2.04</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Machine Tools</td>
<td>54.75</td>
<td>27.00</td>
<td>2.03</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Motor/Generator</td>
<td>46.77</td>
<td>23.69</td>
<td>1.97</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Aerospace</td>
<td>52.18</td>
<td>27.15</td>
<td>1.92</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Chemicals, nec</td>
<td>53.21</td>
<td>28.93</td>
<td>1.84</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Paper</td>
<td>56.94</td>
<td>32.26</td>
<td>1.76</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Other Manufacturing</td>
<td>50.47</td>
<td>28.62</td>
<td>1.76</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Petroleum</td>
<td>59.67</td>
<td>35.00</td>
<td>1.70</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>Glass</td>
<td>52.50</td>
<td>30.83</td>
<td>1.70</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>TV/Radio</td>
<td>48.75</td>
<td>28.75</td>
<td>1.70</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Autoparts</td>
<td>53.64</td>
<td>34.35</td>
<td>1.56</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>Basic Chemicals</td>
<td>53.22</td>
<td>34.29</td>
<td>1.55</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Metal Products</td>
<td>44.63</td>
<td>30.97</td>
<td>1.44</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Drugs</td>
<td>60.85</td>
<td>43.18</td>
<td>1.41</td>
<td>67</td>
<td></td>
</tr>
<tr>
<td>General purpose Machinery, nec</td>
<td>43.37</td>
<td>31.20</td>
<td>1.39</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Electrical Equipment</td>
<td>35.34</td>
<td>26.82</td>
<td>1.32</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>Car/Truck</td>
<td>38.33</td>
<td>30.28</td>
<td>1.27</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Computers</td>
<td>43.35</td>
<td>35.63</td>
<td>1.22</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Medical Equipment</td>
<td>50.11</td>
<td>44.36</td>
<td>1.13</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>Special Purpose Machinery, nec</td>
<td>43.46</td>
<td>38.70</td>
<td>1.12</td>
<td>84</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** Column II is the average between product and process innovations of the mean percentage of innovations for which trade secrecy is considered effective. Column III is the same calculation as Column II for patents. Column IV measures the relative effectiveness of trade secret protection to patent protection.
### Table A1.2  
**Mean and Relative Effectiveness of Intellectual Property Protection Mechanisms**  
**Across Product and Process Innovations by Industry Sector**

<table>
<thead>
<tr>
<th>Industry Sector</th>
<th>((S_{product}+S_{process})/2)</th>
<th>((P_{product}+P_{process})/2)</th>
<th>((S_{product}+S_{process})/(P_{product}+P_{process}))</th>
<th>(N_{firms})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electronics (non-Computer)</td>
<td>46.18</td>
<td>20.62</td>
<td>2.24</td>
<td>48</td>
</tr>
<tr>
<td>Rubber and Plastic</td>
<td>58.00</td>
<td>26.29</td>
<td>2.21</td>
<td>35</td>
</tr>
<tr>
<td>Precision Instruments</td>
<td>45.42</td>
<td>21.32</td>
<td>2.13</td>
<td>18</td>
</tr>
<tr>
<td>Chemicals (non-Petroleum)</td>
<td>59.88</td>
<td>30.94</td>
<td>1.94</td>
<td>176</td>
</tr>
<tr>
<td>Inorganic Materials (non-Metal)</td>
<td>49.44</td>
<td>25.35</td>
<td>1.95</td>
<td>67</td>
</tr>
<tr>
<td>Petroleum</td>
<td>59.67</td>
<td>35.00</td>
<td>1.70</td>
<td>22</td>
</tr>
<tr>
<td>Metal Products (non-Steel)</td>
<td>44.63</td>
<td>30.97</td>
<td>1.44</td>
<td>8</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td>60.85</td>
<td>43.18</td>
<td>1.41</td>
<td>67</td>
</tr>
<tr>
<td>Electrical Equipment</td>
<td>35.34</td>
<td>26.82</td>
<td>1.32</td>
<td>38</td>
</tr>
</tbody>
</table>

**Composition of Industry Sectors**

<table>
<thead>
<tr>
<th></th>
<th>(S_{product}+S_{process})</th>
<th>(P_{product}+P_{process})</th>
<th>((S_{product}+S_{process})/(P_{product}+P_{process}))</th>
<th>(N_{firms})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sectors with High Secrecy to Patent Ratio</td>
<td>50.14</td>
<td>22.71</td>
<td>2.21</td>
<td>101</td>
</tr>
<tr>
<td>Sectors with Average Secrecy to Patent Ratio</td>
<td>57.22</td>
<td>29.86</td>
<td>1.92</td>
<td>265</td>
</tr>
<tr>
<td>Sectors with Low Secrecy to Patent Ratio</td>
<td>51.12</td>
<td>36.81</td>
<td>1.39</td>
<td>113</td>
</tr>
</tbody>
</table>

*Notes: Column II is the average between product and process innovations of the mean percentage of innovations for which trade secrecy is considered effective. Column III is the same calculation as Column II for patents. Column IV measures the relative effectiveness of trade secret protection to patent protection.*

### Table A1.3  
**Total Number of Patent Applications for Seven Years Beginning March 15, 2001 for Mapped Patent Classes in Pennsylvania and New Jersey**

<table>
<thead>
<tr>
<th>Patent Class Arranged by Industry Sector or Composition of Sectors</th>
<th>PA</th>
<th>NJ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Electronics (non-Computer)</strong>*</td>
<td>730</td>
<td>1204</td>
</tr>
<tr>
<td>216 Etching a substrate: processes</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>257 Active solid-state devices (e.g., transistors, solid-state diodes)*</td>
<td>140</td>
<td>164</td>
</tr>
<tr>
<td>326 Electronic digital logic circuitry</td>
<td>6</td>
<td>19</td>
</tr>
<tr>
<td>327 Miscellaneous active electrical nonlinear devices, circuits, and systems</td>
<td>17</td>
<td>28</td>
</tr>
<tr>
<td>329 Demodulators</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>332 Modulators</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>334 Tuners</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>336 Inductor devices</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>343 Communications: radio wave antennas</td>
<td>26</td>
<td>16</td>
</tr>
<tr>
<td>365 Static information storage and retrieval</td>
<td>12</td>
<td>36</td>
</tr>
<tr>
<td>370 Multiplex communications*</td>
<td>122</td>
<td>422</td>
</tr>
<tr>
<td>375 Pulse or digital communications*</td>
<td>118</td>
<td>153</td>
</tr>
<tr>
<td>398 Optical communications*</td>
<td>26</td>
<td>71</td>
</tr>
<tr>
<td>438 Semiconductor device manufacturing: process*</td>
<td>94</td>
<td>115</td>
</tr>
<tr>
<td>455 Telecommunications*</td>
<td>126</td>
<td>217</td>
</tr>
<tr>
<td>720 Dynamic optical information storage or retrieval</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>725 Interactive video distribution systems*</td>
<td>133</td>
<td>88</td>
</tr>
</tbody>
</table>

**Rubber and Plastic**

| 527 Synthetic resins or natural rubbers -- part of the class 520 series | 0 | 3 |

---

96
### Precision Instruments

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
<th>1984</th>
<th>1985</th>
</tr>
</thead>
<tbody>
<tr>
<td>337</td>
<td>Electricity: electrothermally or thermally actuated switches*</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>351</td>
<td>Optics: eye examining, vision testing and correcting</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>352</td>
<td>Optics: motion pictures</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>353</td>
<td>Optics: image projectors</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>355</td>
<td>Photocopying</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>368</td>
<td>Static information storage and retrieval</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>374</td>
<td>Thermal measuring and testing</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>396</td>
<td>Photography</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>399</td>
<td>Electrophotography</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>433</td>
<td>Dentistry</td>
<td>15</td>
<td>6</td>
</tr>
<tr>
<td>602</td>
<td>Surgery: splint, brace, or bandage</td>
<td>7</td>
<td>21</td>
</tr>
<tr>
<td>606</td>
<td>Surgery*</td>
<td>102</td>
<td>255</td>
</tr>
<tr>
<td>623</td>
<td>Prosthesis (i.e., artificial body members), parts thereof, or aids and accessories therefor*</td>
<td>38</td>
<td>174</td>
</tr>
<tr>
<td>702</td>
<td>Data processing: measuring, calibrating, or testing*</td>
<td>55</td>
<td>105</td>
</tr>
</tbody>
</table>

### Chemicals (non-Petroleum)*

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
<th>1984</th>
<th>1985</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>Chemistry: physical processes</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>71</td>
<td>Chemistry: fertilizers</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>95</td>
<td>Gas separation: processes*</td>
<td>23</td>
<td>19</td>
</tr>
<tr>
<td>149</td>
<td>Explosive and thermic compositions or charges</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>201</td>
<td>Distillation: processes, thermolytic</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>260</td>
<td>Chemistry of carbon compounds</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>504</td>
<td>Plant protecting and regulating compositions</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>507</td>
<td>Earth boring, well treating, and oil field chemistry</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>510</td>
<td>Cleaning compositions for solid surfaces, auxiliary compositions therefor, or processes of preparing the compositions</td>
<td>10</td>
<td>84</td>
</tr>
<tr>
<td>512</td>
<td>Perfume compositions</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>516</td>
<td>Colloid systems and wetting agents; subcombinations thereof; processes of</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>518</td>
<td>Chemistry: fischer-tropsch processes; or purification or recovery of products thereof</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>520</td>
<td>Synthetic resins or natural rubbers -- part of the class 520 series</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>522</td>
<td>Synthetic resins or natural rubbers -- part of the class 520 series</td>
<td>6</td>
<td>16</td>
</tr>
<tr>
<td>534</td>
<td>Organic compounds -- part of the class 532-570 series</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>540</td>
<td>Organic compounds -- part of the class 532-570 series*</td>
<td>14</td>
<td>166</td>
</tr>
<tr>
<td>548</td>
<td>Organic compounds -- part of the class 532-570 series*</td>
<td>64</td>
<td>339</td>
</tr>
<tr>
<td>549</td>
<td>Organic compounds -- part of the class 532-570 series*</td>
<td>29</td>
<td>99</td>
</tr>
<tr>
<td>552</td>
<td>Organic compounds -- part of the class 532-570 series</td>
<td>1</td>
<td>11</td>
</tr>
<tr>
<td>554</td>
<td>Organic compounds -- part of the class 532-570 series</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>556</td>
<td>Organic compounds -- part of the class 532-570 series</td>
<td>115</td>
<td>6</td>
</tr>
<tr>
<td>558</td>
<td>Organic compounds -- part of the class 532-570 series</td>
<td>12</td>
<td>38</td>
</tr>
<tr>
<td>570</td>
<td>Organic compounds -- part of the class 532-570 series</td>
<td>2</td>
<td>32</td>
</tr>
<tr>
<td>585</td>
<td>Chemistry of hydrocarbon compounds</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>987</td>
<td>Organic compounds containing a bi, sb, as, or p atom or containing a metal atom of the 6th to 8th group of the periodic system</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Inorganic Materials (non-Metal)

<table>
<thead>
<tr>
<th>Class</th>
<th>Description</th>
<th>1984</th>
<th>1985</th>
</tr>
</thead>
<tbody>
<tr>
<td>51</td>
<td>Abrasive tool making process, material, or composition</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>501</td>
<td>Compositions: ceramic</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>Industry Sector</td>
<td>Patent Classes</td>
<td>Patent Classes</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>----------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>Metal Products (non-Steel)*</td>
<td>43</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>7 Compound tools</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>70 Locks</td>
<td>4</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>75 Specialized metallurgical processes, compositions for use therein, consolidated metal powder compositions, and loose metal particulate mixtures</td>
<td>11</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>109 Safes, bank protection, or a related device</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>232 Deposit and collection receptacles</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>245 Wire fabrics and structure</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>292 Closure fasteners</td>
<td>16</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>419 Powder metallurgy processes</td>
<td>5</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>420 Alloys or metallic compositions</td>
<td>7</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Electrical Equipment*</td>
<td>76</td>
<td>142</td>
<td></td>
</tr>
<tr>
<td>200 Electricity: circuit makers and breakers</td>
<td>6</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>218 High-voltage switches with arc preventing or extinguishing devices</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>320 Electricity: battery or capacitor charging or discharging</td>
<td>6</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>323 Electricity: power supply or regulation systems</td>
<td>6</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>324 Electricity: measuring and testing*</td>
<td>32</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>363 Electric power conversion systems</td>
<td>6</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>373 Industrial electric heating furnaces</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>378 X-ray or gamma ray systems or devices</td>
<td>16</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>601 Surgery: kinesitherapy</td>
<td>3</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Petroleum</td>
<td>6</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>208 Mineral oils: processes and products</td>
<td>6</td>
<td>35</td>
<td></td>
</tr>
</tbody>
</table>

**Compositions of Industry Sectors**

<table>
<thead>
<tr>
<th>Composition</th>
<th>Patent Classes</th>
<th>Patent Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Patent Applications*</td>
<td>5899</td>
<td>10461</td>
</tr>
<tr>
<td>All Patent Applications in Sample*</td>
<td>1281</td>
<td>2802</td>
</tr>
<tr>
<td>Identified Sectors with High Secrecy to Patent Ratio*</td>
<td>962</td>
<td>1775</td>
</tr>
<tr>
<td>Identified Sectors with Average Secrecy to Patent Ratio*</td>
<td>203</td>
<td>837</td>
</tr>
<tr>
<td>Identified Sectors with Low Secrecy to Patent Ratio*</td>
<td>118</td>
<td>190</td>
</tr>
</tbody>
</table>

**Notes:** Industry sectors and compositions of industry sectors are bolded in Column I. Patent classes are listed by number and title in Column I. The sector composed of drugs and medicines is missing because it was not uniquely associated with any patent classes.

* Time series data collected.

<table>
<thead>
<tr>
<th>Patent Class or Set of Classes</th>
<th>State</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Electronics (non-Computer)</strong></td>
<td>PA</td>
<td>50</td>
<td>74</td>
<td>69</td>
<td>84</td>
<td>75</td>
<td>172</td>
<td>204</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>45</td>
<td>85</td>
<td>116</td>
<td>141</td>
<td>240</td>
<td>251</td>
<td>326</td>
</tr>
<tr>
<td>Active solid-state devices (e.g., transistors, solid-state</td>
<td>PA</td>
<td>9</td>
<td>19</td>
<td>13</td>
<td>20</td>
<td>12</td>
<td>31</td>
<td>36</td>
</tr>
<tr>
<td>diodes)</td>
<td>NJ</td>
<td>8</td>
<td>12</td>
<td>29</td>
<td>9</td>
<td>24</td>
<td>34</td>
<td>48</td>
</tr>
<tr>
<td>Multiplex communications</td>
<td>PA</td>
<td>9</td>
<td>17</td>
<td>6</td>
<td>14</td>
<td>16</td>
<td>29</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>12</td>
<td>38</td>
<td>27</td>
<td>53</td>
<td>81</td>
<td>99</td>
<td>132</td>
</tr>
<tr>
<td>Pulse or digital communications</td>
<td>PA</td>
<td>13</td>
<td>17</td>
<td>9</td>
<td>12</td>
<td>9</td>
<td>26</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>11</td>
<td>13</td>
<td>14</td>
<td>18</td>
<td>32</td>
<td>33</td>
<td>32</td>
</tr>
<tr>
<td>Optical communications</td>
<td>PA</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>14</td>
<td>8</td>
<td>38</td>
</tr>
<tr>
<td>Semiconductor device manufacturing: process</td>
<td>PA</td>
<td>7</td>
<td>13</td>
<td>9</td>
<td>19</td>
<td>9</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>4</td>
<td>9</td>
<td>12</td>
<td>15</td>
<td>20</td>
<td>22</td>
<td>33</td>
</tr>
<tr>
<td>Telecommunications</td>
<td>PA</td>
<td>3</td>
<td>8</td>
<td>23</td>
<td>11</td>
<td>17</td>
<td>34</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>4</td>
<td>12</td>
<td>24</td>
<td>29</td>
<td>48</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>Interactive video distribution systems</td>
<td>PA</td>
<td>13</td>
<td>8</td>
<td>6</td>
<td>15</td>
<td>7</td>
<td>37</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>9</td>
<td>6</td>
<td>15</td>
<td>15</td>
<td>9</td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td><strong>Precision Instruments</strong></td>
<td>PA</td>
<td>5</td>
<td>9</td>
<td>12</td>
<td>31</td>
<td>36</td>
<td>54</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>8</td>
<td>27</td>
<td>60</td>
<td>77</td>
<td>97</td>
<td>142</td>
<td>158</td>
</tr>
<tr>
<td>Surgery</td>
<td>PA</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>10</td>
<td>20</td>
<td>24</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>2</td>
<td>15</td>
<td>33</td>
<td>41</td>
<td>36</td>
<td>68</td>
<td>60</td>
</tr>
<tr>
<td>Prosthesis (i.e., artificial body members), parts thereof,</td>
<td>PA</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>or aids and accessories therefor</td>
<td>NJ</td>
<td>0</td>
<td>8</td>
<td>20</td>
<td>25</td>
<td>31</td>
<td>46</td>
<td>44</td>
</tr>
<tr>
<td>Data processing: measuring, calibrating, or testing</td>
<td>PA</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>7</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>5</td>
<td>2</td>
<td>8</td>
<td>10</td>
<td>19</td>
<td>25</td>
<td>36</td>
</tr>
<tr>
<td><strong>Chemicals (non-Petroleum)</strong></td>
<td>PA</td>
<td>9</td>
<td>7</td>
<td>8</td>
<td>20</td>
<td>31</td>
<td>41</td>
<td>68</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>43</td>
<td>86</td>
<td>81</td>
<td>122</td>
<td>133</td>
<td>171</td>
<td>165</td>
</tr>
<tr>
<td>Organic compounds -- part of the class 532-570 series (548)</td>
<td>PA</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>13</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>21</td>
<td>47</td>
<td>41</td>
<td>50</td>
<td>45</td>
<td>67</td>
<td>68</td>
</tr>
<tr>
<td>Organic compounds -- part of the class 532-570 series (549)</td>
<td>PA</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>11</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>4</td>
<td>20</td>
<td>5</td>
<td>16</td>
<td>17</td>
<td>16</td>
<td>21</td>
</tr>
<tr>
<td><strong>Metal Products (non-Steel)</strong></td>
<td>PA</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>8</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>0</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>11</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td><strong>Electrical Equipment</strong></td>
<td>PA</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>9</td>
<td>14</td>
<td>18</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>2</td>
<td>6</td>
<td>14</td>
<td>21</td>
<td>23</td>
<td>34</td>
<td>42</td>
</tr>
<tr>
<td>Electricity: measuring and testing</td>
<td>PA</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>6</td>
<td>9</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>2</td>
<td>5</td>
<td>5</td>
<td>15</td>
<td>13</td>
<td>10</td>
<td>21</td>
</tr>
<tr>
<td><strong>All Patent Applications</strong></td>
<td>PA</td>
<td>260</td>
<td>392</td>
<td>489</td>
<td>752</td>
<td>878</td>
<td>1308</td>
<td>1820</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>431</td>
<td>802</td>
<td>978</td>
<td>1316</td>
<td>1780</td>
<td>2316</td>
<td>2838</td>
</tr>
<tr>
<td><strong>All Patent Applications in Sample</strong></td>
<td>PA</td>
<td>70</td>
<td>98</td>
<td>96</td>
<td>149</td>
<td>166</td>
<td>300</td>
<td>405</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>98</td>
<td>207</td>
<td>278</td>
<td>377</td>
<td>507</td>
<td>615</td>
<td>720</td>
</tr>
<tr>
<td><strong>Identified Sectors with High Secrecy to Patent Ratio</strong></td>
<td>PA</td>
<td>55</td>
<td>83</td>
<td>81</td>
<td>115</td>
<td>111</td>
<td>225</td>
<td>290</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>53</td>
<td>112</td>
<td>176</td>
<td>219</td>
<td>337</td>
<td>393</td>
<td>485</td>
</tr>
<tr>
<td><strong>Identified Sectors with Average Secrecy to Patent Ratio</strong></td>
<td>PA</td>
<td>11</td>
<td>10</td>
<td>11</td>
<td>21</td>
<td>33</td>
<td>45</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>43</td>
<td>87</td>
<td>83</td>
<td>130</td>
<td>136</td>
<td>180</td>
<td>178</td>
</tr>
<tr>
<td><strong>Identified Sectors with Low Secrecy to Patent Ratio</strong></td>
<td>PA</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>13</td>
<td>22</td>
<td>30</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>NJ</td>
<td>2</td>
<td>8</td>
<td>19</td>
<td>28</td>
<td>34</td>
<td>42</td>
<td>57</td>
</tr>
</tbody>
</table>

**Notes:** All sets of patent classes, as represented industry sectors or compositions of sectors, are bolded in Column I. Years begin on March 15 and end on the following March 14. For example, the data in Column I fall between March 15, 2001 and March 14, 2002.
<table>
<thead>
<tr>
<th>State</th>
<th>1999</th>
<th>2000</th>
<th>2001</th>
<th>2002</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pennsylvania</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>46</td>
</tr>
<tr>
<td>New Jersey</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>11</td>
</tr>
</tbody>
</table>

Notes: Years begin on April 19 and end on the following April 18. For example, the data in Column I fall between April 19, 1999 and March 14, 2002.
A1.2 Figures

Figure A1.1  Surjective Mapping of OTAF Sequence Numbers onto Industry Sectors

<table>
<thead>
<tr>
<th>Industry Sector</th>
<th>OTAF Sequence Number and Product Field Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electronics (non-computer)</td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Radio and Television Receiving Equipment Except Communication Types</td>
</tr>
<tr>
<td>43</td>
<td>Electronic Components and Accessories and Communications Equipment</td>
</tr>
<tr>
<td>Rubber and Plastic</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Rubber and Miscellaneous Plastics Products</td>
</tr>
<tr>
<td>Precision Instruments</td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>Professional and Scientific Instruments</td>
</tr>
<tr>
<td>6</td>
<td>Industrial Inorganic Chemistry</td>
</tr>
<tr>
<td>7</td>
<td>Industrial Organic Chemistry</td>
</tr>
<tr>
<td>8</td>
<td>Plastic Materials and Synthetic Resins</td>
</tr>
<tr>
<td>Chemicals (non-Petroleum)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Agricultural Chemicals</td>
</tr>
<tr>
<td>11</td>
<td>Soaps, Detergents, Cleaners, Perfumes, Cosmetics and Toiletries</td>
</tr>
<tr>
<td>12</td>
<td>Paints, Varnishes, Lacquers, Enamels and Allied Products</td>
</tr>
<tr>
<td>13</td>
<td>Miscellaneous Chemical Products</td>
</tr>
<tr>
<td>Inorganic Materials (non-Metal)</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Stone, Clay, Glass and Concrete Products</td>
</tr>
<tr>
<td>Petroleum</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Petroleum and Natural Gas Extraction and Refining</td>
</tr>
<tr>
<td>19</td>
<td>Primary Ferrous Products</td>
</tr>
<tr>
<td>Metal Products (non-Steel)</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Primary and Secondary non-Ferrous Metals</td>
</tr>
<tr>
<td>21</td>
<td>Fabricated Metal Products</td>
</tr>
<tr>
<td>Pharmaceuticals</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Drugs and Medicines</td>
</tr>
<tr>
<td>Electrical Equipment</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>Electrical Transmission and Distribution Equipment</td>
</tr>
<tr>
<td>36</td>
<td>Electrical Industrial Apparatus</td>
</tr>
<tr>
<td>37</td>
<td>Household Appliances</td>
</tr>
<tr>
<td>38</td>
<td>Electrical Lighting and Wiring Equipment</td>
</tr>
<tr>
<td>40</td>
<td>Miscellaneous Electrical Machinery, Equipment and Supplies</td>
</tr>
</tbody>
</table>
A2.1 Mapping Patent Class to Industry Sector

clear all;
close all;

% Imports Excel files containing patent application data
c onc2005=csvread('conc2005.csv');

% Creates mapping matrix shells
MAPshell=zeros(1000,57);
CONshell=zeros(1000,10);

% Sets OTAF sequence numbers for desired industries
chemOTAF=[3 4 5 6 7 8 9 10 11 12 13];
drugOTAF=[14];
petrOTAF=[15];
rubbOTAF=[16];
limeOTAF=[17];
metlOTAF=[18 19 20 21];
elecOTAF=[33 34 35 36 37 38 39 40];
ecomOTAF=[41 42 43];
instOTAF=[55];
otherOTAF=[1 2 22 24 25 26 27 28 29 30 31 32 44 45 46 47 48 49 50 51 52 53 54 56 57];

OTAF={otherOTAF, chemOTAF, drugOTAF, petrOTAF, rubbOTAF, limeOTAF, metlOTAF, elecOTAF, ecomOTAF, instOTAF};

% Mapping procedure
for i=1:length(conc2005(:,1));
    MAPshell(conc2005(i,1),conc2005(i,2))=MAPshell(conc2005(i,1),conc2005(i,2))+1;
end

% Condensing procedure
for i=1:length(OTAF);
    for j=1:length(OTAF{i});
        CONshell(:,i)=CONshell(:,i)+MAPshell(:,OTAF{i}{j});
    end
end

% Initialize mapping matrix
MAP=CONshell;

% Elimination procedure (surjective step)
for i=1:length(CONshell(:,1));
    if CONshell(i,1) ~= 0;
        MAP(i,:) = 0*MAP(i,:);
    end
    if length(find(MAP(i,:))) > 1;
        MAP(i,:) = 0;
    end
end

% Removes "other" column from MAP
MAP(:,1) = [];

% Initializes industry-CIL surjective buckets
chemCCL = [];
drugCCL = [];
petrCCL = [];
rubbCCL = [];
limeCCL = [];
metlCCL = [];
elecCCL = [];
ecomCCL = [];
instCCL = [];

CCL = {chemCCL, drugCCL, petrCCL, rubbCCL, limeCCL, metlCCL, elecCCL, ecomCCL, instCCL};

% Data reorganization and readout
for j=1:length(CCL);
    CCL{j} = find(MAP(:,j));
end

chemCCL = CCL{1}
drugCCL = CCL{2}
petrCCL = CCL{3}
rubbCCL = CCL{4}
limeCCL = CCL{5}
metlCCL = CCL{6}
elecCCL = CCL{7}
ecomCCL = CCL{8}
instCCL = CCL{9}

% Created by Mackey Craven on August 1, 2010
A2.2 Performing Difference in Difference Regressions

```matlab
close all;
clear all;

%% IMPORT DATA and SEPERATE OBSERVATIONS

% Imports Excel files containing patent application data
PatentNJImport=xlsread('NJ|patentapp, Matlab.xls');
PatentPAImport=xlsread('PA|patentapp, Matlab.xls');
TradeSecretImport=xlsread('Trade Secret Data.xls');

% Isolates time series data
NJall=PatentNJImport(1:1, 1:7); % All patent applications in the state of NJ
NJor=PatentNJImport(88:88, 1:7); % All patent applications in the NJ sample
NJts=TradeSecretImport(4:4, 1:10); % All trade secret cases in the NJ sample

NJinst1=PatentNJImport(31:31, 1:7); % All patent applications in the NJ sample in class 606
NJinst2=PatentNJImport(32:32, 1:7); % All patent applications in the NJ sample in class 623
NJinst3=PatentNJImport(33:33, 1:7); % All patent applications in the NJ sample in class 702
NJinstor=PatentNJImport(34:34, 1:7); % All patent applications in the NJ sample associated exclusively with the precision instruments industry

NJecom1=PatentNJImport(3:3, 1:7); % All patent applications in the NJ sample in class 257
NJecom2=PatentNJImport(12:12, 1:7); % All patent applications in the NJ sample in class 370
NJecom3=PatentNJImport(13:13, 1:7); % All patent applications in the NJ sample in class 375
NJecom4=PatentNJImport(14:14, 1:7); % All patent applications in the NJ sample in class 398
NJecom5=PatentNJImport(15:15, 1:7); % All patent applications in the NJ sample in class 438
NJecom6=PatentNJImport(16:16, 1:7); % All patent applications in the NJ sample in class 455
NJecom7=PatentNJImport(18:18, 1:7); % All patent applications in the NJ sample in class 725
NJecomor=PatentNJImport(19:19, 1:7); % All patent applications in the NJ sample associated exclusively with the electrical components

NJchem1=PatentNJImport(74:74, 1:7); % All patent applications in the NJ sample in class 548
```
NJchem2=PatentNJImport(75:75, 1:7); % All patent applications in the NJ sample in class 549
NJchemor=PatentNJImport(83:83, 1:7); % All patent applications in the NJ sample associated exclusively with the chemical industry
NJelec1=PatentNJImport(41:41, 1:7); % All patent applications in the NJ sample in class 324
NJelecor=PatentNJImport(46:46, 1:7); % All patent applications in the NJ sample associated exclusively with the electrical equipment industry
NJmetlor=PatentNJImport(56:56, 1:7); % All patent applications in the NJ sample associated exclusively with low ratio industries
NJhighor=PatentNJImport(36:36, 1:7); % All patent applications in the NJ sample associated exclusively with high secrecy ratio industries
NJmidlor=PatentNJImport(87:87, 1:7); % All patent applications in the NJ sample associated exclusively with middle ratio industries
NJlowor=PatentNJImport(57:57, 1:7); % All patent applications in the NJ sample associated exclusively with low ratio industries

PAall=PatentPAImport(1:1, 1:7); % All patent applications in the state of PA
PAor=PatentPAImport(88:88, 1:7); % All patent applications in the PA sample
PAts=TradeSecretImport(4:4, 1:10); % All trade secret cases in the PA sample
PAinst1=PatentPAImport(31:31, 1:7); % All patent applications in the PA sample in class 606
PAinst2=PatentPAImport(32:32, 1:7); % All patent applications in the PA sample in class 623
PAinst3=PatentPAImport(33:33, 1:7); % All patent applications in the PA sample in class 702
PAinstor=PatentPAImport(34:34, 1:7); % All patent applications in the PA sample associated exclusively with the precision instruments industry
PAecom1=PatentPAImport(3:3, 1:7); % All patent applications in the PA sample in class 257
PAecom2=PatentPAImport(12:12, 1:7); % All patent applications in the PA sample in class 370
PAecom3=PatentPAImport(13:13, 1:7); % All patent applications in the PA sample in class 375
PAecom4=PatentPAImport(14:14, 1:7); % All patent applications in the PA sample in class 398
PAecom5=PatentPAImport(15:15, 1:7); % All patent applications in the PA sample in class 438
PAecom6=PatentPAImport(16:16, 1:7); % All patent applications in the PA sample in class 455
PAecom7 = PatentPAImport(18:18, 1:7); % All patent applications in the PA sample in class 725
PAecomor = PatentPAImport(19:19, 1:7); % All patent applications in the PA sample associated exclusively with the electrical components

PAchem1 = PatentPAImport(74:74, 1:7); % All patent applications in the PA sample in class 548
PAchem2 = PatentPAImport(75:75, 1:7); % All patent applications in the PA sample in class 549
PAchemor = PatentPAImport(83:83, 1:7); % All patent applications in the PA sample associated exclusively with the chemical industry

PAelec1 = PatentPAImport(41:41, 1:7); % All patent applications in the PA sample in class 324
PAelecor = PatentPAImport(46:46, 1:7); % All patent applications in the PA sample associated exclusively with the electrical equipment industry

PAmetlor = PatentPAImport(56:56, 1:7); % All patent applications in the PA sample associated exclusively with low ratio industries

PAhighor = PatentPAImport(36:36, 1:7); % All patent applications in the PA sample associated exclusively with high secrecy ratio industries
PAmidlor = PatentPAImport(87:87, 1:7); % All patent applications in the PA sample associated exclusively with middle ratio industries
PALowor = PatentPAImport(57:57, 1:7); % All patent applications in the PA sample associated exclusively with low ratio industries

% Combines NJ and PA observations of the same type into one observation 'Ytype' as a column vector
Yall = vertcat(NJall', PAall');
Yor = vertcat(NJor', PAor');
Yts = vertcat(NJts', PAts');

Yinst1 = vertcat(NJinst1', PAinst1');
Yinst2 = vertcat(NJinst2', PAinst2');
Yinst3 = vertcat(NJinst3', PAinst3');
Yinstor = vertcat(NJinstor', PAinstor');

Yecom1 = vertcat(NJecom1', PAecom1');
Yecom2 = vertcat(NJecom2', PAecom2');
Yecom3 = vertcat(NJecom3', PAecom3');
Yecom4 = vertcat(NJecom4', PAecom4');
Yecom5 = vertcat(NJecom5', PAecom5');
Yecom6 = vertcat(NJecom6', PAecom6');
Yecom7 = vertcat(NJecom7', PAecom7');
Yecomor = vertcat(NJecomor', PAecomor');
Ychem1=vertcat(NJchem1', PAchem1');
Ychem2=vertcat(NJchem2', PAchem2');
Ychemor=vertcat(NJchemor', PAchemor');
Yelec1=vertcat(NJelec1', PAelec1');
Yelecor=vertcat(NJelecor', PAelecor');
Ymetlor=vertcat(NJmetlor', PAmetlor');
Yhighor=vertcat(NJhighor', PAhighor');
Ymidlor=vertcat(NJmidlor', PAmidlor');
Ylowor=vertcat(NJlowor', PAlowor');

%% CREATE DESIGN MATRICES

% Creates shell for X's and sets iteration parameters
Designall=zeros(length(Yall(:,1)),5);
Designor=zeros(length(Yor(:,1)),5);
Designts=zeros(length(Yts(:,1)),5);
Designinst1=zeros(length(Yinst1(:,1)),5);
Designinst2=zeros(length(Yinst2(:,1)),5);
Designinst3=zeros(length(Yinst3(:,1)),5);
Designinstor=zeros(length(Yinstor(:,1)),5);
Designecom1=zeros(length(Yecom1(:,1)),5);
Designecom2=zeros(length(Yecom2(:,1)),5);
Designecom3=zeros(length(Yecom3(:,1)),5);
Designecom4=zeros(length(Yecom4(:,1)),5);
Designecom5=zeros(length(Yecom5(:,1)),5);
Designecom6=zeros(length(Yecom6(:,1)),5);
Designecom7=zeros(length(Yecom7(:,1)),5);
Designecomor=zeros(length(Yecomor(:,1)),5);
Designchem1=zeros(length(Ychem(:,1)),5);
Designchem2=zeros(length(Ychem2(:,1)),5);
Designchemor=zeros(length(Ychemor(:,1)),5);
Designelec1=zeros(length(Yelec1(:,1)),5);
Designelecor=zeros(length(Yelecor(:,1)),5);
Designmetlor=zeros(length(Ymetlor(:,1)),5);
Designhighor=zeros(length(Yhighor(:,1)),5);
Designmidlor=zeros(length(Ymidlor(:,1)),5);
Designlowor=zeros(length(Ylowor(:,1)),5);
Designlowor=zeros(length(Ylowor(:,1)),5);

J=7;
l=length(Designtall(:,1))/J;

% Iteration to fill in design shells

for i=1:l;
    for j=1:J;
        if i<=l/2;
            s=0;  % Signifying control state
            if j<=2;
                e=0;  % Signifying no event in the first two periods
            else
                e=1;  % Signifying event has taken place
            end
        else
            s=1;  % Signifying affected state
            if j<=2;
                e=0;  % Signifying no event in the first two periods
            else
                e=1;  % Signifying event has taken place
            end
        end
    end
Designall(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];  % Maps iteration to row
Designor(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];

Designinst1(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designinst2(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designinst3(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designinstor(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];

Designecom1(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designecom2(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designecom3(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designecom4(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designecom5(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designecom6(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designecom7(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designecomor(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];

Designchem1(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designchem2(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designchemor(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];

Designelec1(J*(i-1)+j,:)=\[1 s e s (j-1) (j-1)*e*s\];
Designelecor(J*(i-1)+j,:)=[1 s e*s (j-1) (j-1)*e*s];

Designmetlor(J*(i-1)+j,:)=[1 s e*s (j-1) (j-1)*e*s];

Designhigheor(J*(i-1)+j,:)=[1 s e*s (j-1) (j-1)*e*s];
Designmidlor(J*(i-1)+j,:)=[1 s e*s (j-1) (j-1)*e*s];
Designlowor(J*(i-1)+j,:)=[1 s e*s (j-1) (j-1)*e*s];
end
end

% Iteration to fill in design shells for secrecy
Jts=10;
lts=length(Designts(:,1))/Jts;
for i=1:lts;
    for j=1:Jts;
        if i<=lts/2;
            s=0; % Signifying control state
            if j<=5;
                e=0; % Signifying no event in the first two periods
            else
                e=1; % Signifying event has taken place
            end
        else
            s=1; % Signifying affected state
            if j<=5;
                e=0; % Signifying no event in the first two periods
            else
                e=1; % Signifying event has taken place
            end
        end
        Designts(Jts*(i-1)+j,:)= [1 s e*s (j-1) (j-1)*e*s]; % Maps iteration to row
    end
end

% Initialization of X matrices
Xall=Designall;
Xor=Designor;
Xts=Designts;

Xinst1=Designinst1;
Xinst2=Designinst2;
Xinst3=Designinst3;
Xinstor=Designinstor;
Xinsttot=Designinsttot;
Xecom1=Designecom1;
Xecom2=Designecom2;
Xecom3=Designecom3;
Xecom4=Designecom4;
Xecom5=Designecom5;
Xecom6=Designecom6;
Xecom7=Designecom7;
Xecomor=Designecomor;
Xcommtot=Designecommtot;

Xchem1=Designchem1;
Xchem2=Designchem2;
Xchemor=Designchemor;
Xchemtot=Designchemtot;

Xelec1=Designelec1;
Xelecor=Designelecor;

Xmetlor=Designmetlor;

Xhighor=Designhighor;
Xmidlor=Designmidlor;
Xlowor=Designlowor;

%% Perform Regressions

Statsall=regstats(Yall,Xall(1:length(Xall(:,1)),
2:length(Xall(:,1))),'linear',{'beta','r','mse','rsquare','tstat','dwstat'});

Statsor=regstats(Yor,Xor(1:length(Xor(:,1)),
2:length(Xor(:,1))),'linear',{'beta','r','mse','rsquare','tstat','dwstat'});

Stats=regstats(Yts,Xts(1:length(Xts(:,1)),
2:length(Xts(:,1))),'linear',{'beta','r','mse','rsquare','tstat','dwstat'});

Statsinst1=regstats(Yinst1,Xinst1(1:length(Xinst1(:,1)),
2:length(Xinst1(:,1))),'linear',{'beta','r','mse','rsquare','tstat','dwstat'});

Statsinst2=regstats(Yinst2,Xinst2(1:length(Xinst2(:,1)),
2:length(Xinst2(:,1))),'linear',{'beta','r','mse','rsquare','tstat','dwstat'});

Statsinst3=regstats(Yinst3,Xinst3(1:length(Xinst3(:,1)),
2:length(Xinst3(:,1))),'linear',{'beta','r','mse','rsquare','tstat','dwstat'});

Statsinstor=regstats(Yinstor,Xinstor(1:length(Xinstor(:,1)),
2:length(Xinstor(:,1))),'linear',{'beta','r','mse','rsquare','tstat','dwstat'});

Statsecom1=regstats(Yecom1,Xecom1(1:length(Xecom1(:,1)),
2:length(Xecom1(:,1))),'linear',{'beta','r','mse','rsquare','tstat','dwstat'});

Statsecom2=regstats(Yecom2,Xecom2(1:length(Xecom2(:,1)),
2:length(Xecom2(:,1))),'linear',{'beta','r','mse','rsquare','tstat','dwstat'});
2:length(Xecom2(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});
Statsecom3=regstats(Yecom3,Xecom3(1:length(Xecom3(:,1)),
2:length(Xecom3(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});
Statsecom4=regstats(Yecom4,Xecom4(1:length(Xecom4(:,1)),
2:length(Xecom4(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});
Statsecom5=regstats(Yecom5,Xecom5(1:length(Xecom5(:,1)),
2:length(Xecom5(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});
Statsecom6=regstats(Yecom6,Xecom6(1:length(Xecom6(:,1)),
2:length(Xecom6(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});
Statsecomor=regstats(Yecomor,Xecomor(1:length(Xecomor(:,1)),
2:length(Xecomor(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});

Statschem1=regstats(Ychem1,Xchem1(1:length(Xchem1(:,1)),
2:length(Xchem1(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});
Statschem2=regstats(Ychem2,Xchem2(1:length(Xchem2(:,1)),
2:length(Xchem2(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});
Statschemor=regstats(Ychemor,Xchemor(1:length(Xchemor(:,1)),
2:length(Xchemor(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});

Statselect=regstats(Yelec1,Xelec1(1:length(Xelec1(:,1)),
2:length(Xelec1(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});
Statselector=regstats(Yelecor,Xelecor(1:length(Xelecor(:,1)),
2:length(Xelecor(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});

Statsmetlor=regstats(Ymetlor,Xmetlor(1:length(Xmetlor(:,1)),
2:length(Xmetlor(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});

Statshighor=regstats(Yhighor,Xhighor(1:length(Xhighor(:,1)),
2:length(Xhighor(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});
Statsmidlor=regstats(Ymidlor,Xmidlor(1:length(Xmidlor(:,1)),
2:length(Xmidlor(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});
Statslowor=regstats(Ylowor,Xlowor(1:length(Xlowor(:,1)),
2:length(Xlowor(1,:))),'linear',
{\text{'beta','r','mse','rsquare','tstat','dwstat'}});

% Created by Mackey Craven on August 3, 2010
Bibliography


E.I. du Pont de Nemours Powder Co. v. Masland, 244 U.S. 100, 102 (1917).


Restatement (First) of Torts § 757 (1939).

Restatement (Second) of Torts (1978).


U.S. Const. art. I, § 8, cl. 8.


18 P.C.S. § 5302 (2010).