UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

Publicly Funded Renewable-Energy Innovation and Appropriate Methods of Analysis

Permalink

https://escholarship.org/uc/item/2wh6514g

Author

Bush, Nathaniel

Publication Date

2017

Peer reviewed|Thesis/dissertation

**Publicly Funded Renewable-Energy Innovation
and Appropriate Methods of Analysis**

By

Nathaniel Bush

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Public Policy

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michael O'Hare, Chair
Professor Eugene Smolensky
Professor Karlene Roberts

Spring 2017

Abstract

Publicly Funded Renewable-Energy Innovation and Appropriate Methods of Analysis

by

Nathaniel Bush

Doctor of Philosophy in Public Policy

University of California, Berkeley

Professor Michael O'Hare, Chair

In this dissertation, I will conduct a critical analysis of several methods typically used for modeling patent data, and then use insights gleaned from that analysis to explore four instances of public financing of innovation in the form of renewable-energy technology.

Chapter 1 introduces my overall research objectives, frames the research in terms of its policy relevance, and provides a brief preview of the major results. Chapter 2 provides background on innovation, causes and results thereof, with a particular focus on patent data, and thereby a framework for understanding the relevance of chapters 3 and 4.

The first main essay (chapter 3) addresses statistical regression techniques frequently used with patent-count time-series data, namely negative binomial regressions and log-log regressions. It reveals high rates of spurious correlation (false positives, also known as Type I errors) when using these techniques on patent data, investigates possible ways of addressing this problem, and creates a method for detecting it when using those and other regression techniques on similar data types.

Using lessons learned from the first essay, the second essay (chapter 4) examines public research and development (R&D) funding of four renewable-energy technologies – wind turbines, solar photovoltaics, solar thermal electric, and solar water heating. It employs several novel patent analysis techniques that allow patents to more closely represent the date of inventive activity. It finds that, contrary to popular narratives of public R&D funding driving increased invention (patenting), a diverse set of relationships exist in the renewable energy sector. These relationships range from changes in funding being correlated with future changes in invention rates, to changes in invention rates preceding changes in funding, to changes in funding and invention rates having no discernible relationship.

Taken together, these essays demonstrate the interdependent relationship between appropriate analytic techniques and accurate analysis when examining the sometimes subtle effects of complex policies.

Dedication

I dedicate this dissertation to my family: my wife, Natalie; my mother, Vickie; and my daughter, Naomi. Each of you has been vital to the completion of this work in your own ways. Without your love, support, understanding, tolerance, and prodding I certainly wouldn't have made it. You have my love. You are my life.

Acknowledgments

I'd like to thank the many people who have contributed their time, energy, knowledge, and encouragement to this work:

To my wife – thank you for not letting my doubts sink me, for providing a great reality check, for supporting us while I wrote this, and most of all for your love.

To my mother – thank you for encouraging and believing in me, and for all of the quality child care you've provided while I worked. You knew I could do it even when I didn't.

To my daughter – thank you for being born while I was working on this! Not only were you a much wanted and welcome addition to our lives, but you provided me with a powerful incentive to finish up.

To my advisor, Michael O'Hare – thank you for taking me on after my first two advisors left the university, and seeing me through to completion. You're the kind of sage professor I had in mind when I imagined going to graduate school, and the kind far more professors should actually be.

To my other two committee members, Geno Smolensky and Karlene Roberts – thank you for all of your advice, wisdom, and encouragement, which started long before I began writing. You are two of the pillars holding up UC Berkeley.

To Cecille Cabacungan – thank you for your considerable efforts in this last year. Your regular communication, timely information, and obvious dedication were unbelievably helpful. I honestly don't think I would have finished if it weren't for you.

To Janet Sawin – thank you for sharing your (amazingly extensive) dissertation on wind power programs, it was enlightening and provided a few missing data points that helped greatly.

To the people I interviewed, Joseph Desmond, Daniel Purdy, Pieter Mul, and Meghan Burton – thank you for your time and insights. You added depth and nuance to what might have been only dry data analysis.

To my dear friends Bob, Ethan, and Peter – thank you for your encouragement and friendship, and for not asking too many times, "Are you done yet?"

Table of Contents

List of Graphs

List of Tables

1   Introduction & Overview

1.1   Overall Research Objectives

Through this dissertation I intend to: (1) explore the efficacy of public funding policies designed to foster innovation in young or emerging technologies, specifically in four branches of renewable-energy technology; (2) utilize patent data as a primary metric of inventive activity in those technology areas; (3) critically examine statistical techniques typically applied to patent-data analysis, and thereby bring to bear an improved set of methods in my own analyses; and, (4) elucidate previously undetected relationships between funding and invention to generate policy-relevant information.

1.2   Policy Relevance

As a largely capitalist society, many people tend to take the perspective that firms should innovate as people's demands change, as a function of their profit motive and desire to gain customers. From this perspective, with innovation as a method of gaining market share and profit, there is no obvious purpose for government intervention. However, when viewed from a societal level, the beneficial spillovers (i.e. positive externalities and technological complementarities) that innovation generates, to Gross Domestic Product (GDP) and other measures of wellbeing (Bresnahan and Trajtenberg 1995, Fagerberg and Verspagen 2002, Greiner and Semmler 2000, and Carlaw and Lipsey 2002), mean that we need to consider the fruits of innovation as being policy relevant. Moreover, as we are faced with the effects of negative externalities (e.g. climate change) that we hope innovation will ameliorate without the need to drastically alter our civilization, it is logical to ask whether current incentive structures will foster the needed technological development and diffusion. If we do not believe that purely free-market incentives will do so independently, then we need to cultivate policies that will generate the needed change. Innovation becomes an issue of concern to public policy when viewed from the proper altitude.

Young and emerging technologies are of particular interest to policy makers and analysts due to their capacity to disrupt existing modes of living, produce novel and/or outsized benefits compared to their costs, generate unforeseen consequences and damage, and create unpredictable spillover effects (Giersch 1982, Carlaw and Lipsey 2002). My dissertation will focus on four young renewable-energy technologies with particular interest to public policy: wind turbines, solar photovoltaics, solar thermal electric generation, and solar water heating. Researchers advancing all of these technology areas tout their potential to mitigate existing environmental externalities, while critics claim they have the capacity to generate a host of new problems. In order to inform applicable policy decision making, it will be useful to understand both the patterns of development of these technologies and how research activity has responded to existing incentives.

In addition to examining particular technologies, my dissertation will generate higher quality policy-relevant information by improving the set of methods used to conduct patent-data analysis. Not only will I employ better selection techniques for the inclusion of patents than prior research, and deal with issues of pendency and priority date more accurately than much of the

existing research, but I will test two statistical methods frequently applied to patent-data research for possible spurious results. In this manner, I hope to conduct higher quality policy-relevant research and help inform other researchers' efforts to do the same.

Given that my methods-related findings inform the techniques I use on the renewable-energy technology section, I will present the methods work in chapter 3 and the renewable-energy patent analysis in chapter 4.

1.3   Results in Brief

Chapter 3 of this dissertation presents results from Monte Carlo simulations of negative binomial and log-log regressions using 10,000 synthetic data series that mimic the statistical properties of patent data. These simulations yielded rates of spurious correlations well in excess of the expected 5 percent level; both falling in the 60 to 80 percent range. This is due to non-stationarity (i.e. no fixed mean, causing errors to accumulate) in patent data, and this indicates the need to test and account for this problem when dealing with patent data. Independent variables include ones commonly used in patent-data research, such as US Gross Domestic Product (USGDP) and a composite price series.

Scholars who analyze patent data frequently employ two non-traditional regression techniques – negative-binomial, which models the dependent variable as an integer count truncated at zero, and log-log, which log transforms both the independent and dependent variables before performing an ordinary least squares (OLS) regression – but the validity of these techniques is seriously called into question by the results presented. These techniques were intended to better represent the data generating process and avoid the problems associated with using OLS, but end up generating comparable levels of spurious correlation to OLS. This in turn calls into question the results of a substantial number of papers in the innovation literature.

The results also demonstrate the capacity to generate synthetic time-series to accurately mimic the behavior of patent-data, and use these to test the validity of statistical methods. With false positives likely in the range of 60 to 80 percent, the inability to rule out a "random walk" (where the errors from all prior time periods accumulate) as the data generating process means we cannot trust results from these regression methods when used with non-stationary independent variables on patent data. It also may explain the popularity of some of these techniques, since using them has a high probability of generating spurious "statistically significant" results.

Finally, these results demonstrate the potential for Monte Carlo simulations to be used as a validating method when concerns about the legitimacy of a more exotic statistical technique arise.

Using the lessons learned in chapter 3, chapter 4 of this dissertation examines the relationships between federal government R&D expenditure and renewable-energy technology patenting through 20 Vector Autoregression (VAR) models. Patenting rates for four renewable energy technologies (and a pooled patent series) are used, with models for all patenting and for government-interest-only patenting.

Diverse modeling outcomes result – from R&D spending being Granger-caused by patenting (this is a weak causal claim originally proposed by Clive Granger [1969], where past values of one variable are correlated to present values of another variable), to patents being

Granger-caused by R&D spending, to patents and R&D spending Granger-causing each other. These results indicate that policy design details are vitally important in the case of government funding for renewable energy R&D – merely "throwing money at the problem" does not guarantee the desired result. This is also suggested by the changes in modeling results between the all-patent models and the government-interest models – the isolation of government-interest patents reveals a more pervasive situation of patenting Granger-causing R&D funding than was seen in the all-patents models. Incentive structures, institutional relationships, geography, and methods of knowledge sharing within a given funding program may determine if public R&D funding is successful in actually promoting targeted invention, or merely responds to newly popular technology areas.

Scholarly articles noted in chapter 4 overlook complex policy details and nuances when discussing public research funding, and my analysis demonstrates the importance of attending to such details, since the relationships of patenting rates and public R&D funding differ widely depending on the technology and funding program.

For the purposes of motivating near-term affordability of functioning technologies – ones that already have some demonstrated capabilities, but are not yet cost competitive with conventional energy sources – different approaches are needed. Public funding for commercialization oriented renewable-energy invention seems to be a less certain method; at least with the types of funding programs that have existed so far.

In the next section, chapter 2, I will provide relevant background on innovation, and on patent-data research in particular.

2　　Background on Innovation

In this chapter, I will explore the study of innovation, looking at models and measures from economics, organizational/management theory, and public policy. I begin by defining innovation and looking at its origins as an area of academic study, starting with Schumpeter. I then describe identified categories of innovation, factors influencing innovation, and its theoretical outcomes. Next, I focus on proxies used to measure innovation, including production shifts, patents and citations, publications (i.e. academic research papers and reports), intra-firm measures of innovation (i.e. counting knowledge workers, research projects, research expenditures, and revenue from recent innovation), and other measures (e.g. impact on environmental externalities and technology balance of payments). Finally, I explore why innovation is an issue of concern to public policy, and discuss the various policy instruments used to influence innovation – as well as their shortcomings.

2.1　　What Is Innovation, and How Do We Measure and Affect It?

All equilibrium narratives end in the same place. In equilibrium. There's no role for a hero, tragic or other. Even the true protagonist, the Walrasian Auctioneer, is hidden away, unexplained. Growth stories, in contrast, are more open-ended, uncertain, and path-dependent. In growth stories, history matters. Even tragedies, which also end predictably, must tell the beginning first. Thus do political economy's best-known growth stories provide a hero. (Leonard 2009)

The ultimate hero of the modern world is innovation. Innovation has literally made the difference between humans wandering around naked, eating whatever we find, and living comfortably in modern industrial society. In that way, it is one of the ultimate generators of disequilibrium in the world, as well as path dependence, and a host of other factors that complicate simple economic analysis.

I will broadly define innovation as the creation and propagation of new ideas. While authors have created many definitions of innovation, creation and propagation are at the core of each. Importantly, innovation does not stop at creation. The distinction Schumpeter (1942) makes between invention and innovation is vital. "An invention is an idea, sketch, or model for a new device, process or system. It might be patented or not, it might lead to innovation or not" (Clarke and Riba 1998). Innovation requires something more: "adoption" according to Schumpeter, which is the initial use of an invention, and "diffusion," which is broader commercialization due to operating information and user communication (Rogers 1995). Since inventions may be diffused very narrowly (e.g. an idiosyncratic change to a type of power-plant's fuel mix) or more broadly (e.g. the integrated circuit), I consider adoption and diffusion together as elements of propagation.

Although innovation may be simple to define in general, highly varied patterns of it exist: it may occur radically or incrementally, it may create new systems or affect existing ones, it may have broad spillovers or highly specific impacts, and it may be geographically localized or global in reach. Innovation, and thereby change, was a basic driver of capitalism according to Schumpeter. He wrote, "[Capitalism] not only never is but never can be stationary . . ." also:

*[I]n capitalist reality as distinguished from its textbook picture, it is not...
[textbook]competition which counts but the competition from the new commodity,
the new technology, the new source of supply, the new type of organization (the
largest-scale unit of control for instance) – competition which commands a
decisive cost or quality advantage and which strikes not at the margins of the
profits and the outputs of the existing firms but at their foundations and their very
lives.* (Schumpeter 1942)

In order to study innovation (or anything else for that matter), it must be measured. More
than just a practical problem, the question of what to measure when examining innovation is a
philosophical issue: insofar as innovation induces change and evolution in systems, the measures
that were developed to suit one incarnation of a system may not remain totally applicable to
another; the definitions of progress and change in a system may themselves change, yielding
non-comparability between past and present measures, results, and frames of reference.

Traditional measures of innovation, such as productivity growth and patents, have been
augmented with additional measures, such as academic publications and citations, various firm-
level characteristics, international balance of payments, and environmental impacts as the
economy has become more complicated, diverse, and global.

Policy instruments to influence innovation are similarly diverse, and include taxes,
subsidies, permits, intellectual property laws, public research funding, and "command-and-
control" regulations, among others. Each instrument has a host of complexities: conditions under
which it is counter-productive, problems for which it is and is not suitable, indirect and/or
unpredictable impacts, and ways that it may be manipulated.

2.2   Innovation Inputs and Outputs

Economics and organizational/management theory have generated several models for
thinking about innovation. These models pertain to various types of innovation that can occur,
the sources and causes of innovation, and the results of innovation. In this section, I will examine
several of these factors, with an eye to not only exploring how academics have viewed
innovation, but also to building up a framework that can be used to understand how the measures
of innovation succeed and fail.

2.2.1   Determinants of Innovation

Different forms of innovation may result from different social forces, but given their
common framework, they are likely to have consistent themes. Product innovations, service
innovations, and organizational innovations may also vary in their causal sources, however,
insofar as they all relate to the desire to do something new or make an existing thing better, they
also have common threads. I will consider various theoretical determinants of innovation, and
what forms of innovation they relate to.

While the profit motive, or some incarnation of it (e.g. prestige or intellectual satisfaction
in academic research), is the fundamental incentive to innovate, many other factors mold how
that incentive is expressed in the world, and what causes the particular patterns of innovation that

occur.

The shaping of innovation is an area of study extremely rich in modeling work. Given social-science's ultimate focus on prediction, this is to be expected. Various types of determinants have been examined, including intellectual property rights, government actions/regulation, knowledge stock, geography, management and organizational characteristics, and influences from and on academic research.

2.2.1.1   Intellectual Property Rights

Intellectual property (IP) rights are actually a subset of government actions; however, they are so vital to inducing innovation that I have broken them out into their own section here. Common forms of IP relevant to innovation are patents, copyrights, and trade secrets. Their existence enables individuals or firms to maintain profitable control over inventions, and society and markets to signal the value of various inventions and the entities that own them (by means of exchanging information for protection) (Long 2002).

Possibly the most important factor influencing the incentive to innovate, and thereby actual innovative activity, is an entity's ability to benefit from that innovation. In this regard, the existence of IP rights is a central determinant of innovation. However, "some of the 'ineffectiveness' of formal intellectual property protection is by design… it is not always true that stronger intellectual property rights are better. Intellectual property should be designed to achieve the right balance of protection for innovators, protection for consumers, and opportunity for rivals to make improvements" (Scotchmer 2004b). By providing protections strong enough to allow innovators to profit, but not so strong as to stifle competition, IP rights attempt to optimize creative incentives.

While the general existence (and strength) of IP rights is an input into innovation, individual pieces of IP are rightly viewed as products of innovative activity (e.g. patents are sought after an invention has been conceived of, and copyrights are taken out on existing works). Therefore, I discuss particular types of IP as measures of innovation in sub-sections of section 2.2.3 below.

2.2.1.2   Government Actions/Regulation

A variety of government actions can stimulate technological innovation, including the provision of positive inducements, such as tax breaks, subsidies, contracts, and prizes; the facilitation of knowledge sharing; and the outright prohibition of certain behaviors, substances, or products. Insofar as government actions and laws represent the constraints and restrictions under which markets operate, they are also foundational to most other determinants of innovation, especially the above mentioned intellectual property rights. Since public academic research is discussed below, and positive government inducements to innovate constitute little more than a shift in the probability distribution of profits accruing to innovations, I will focus primarily on regulation here.

A study of multiple factors influencing environmental innovation found that "government regulation appears to be a greater stimulus to inventive activity than government-sponsored research support alone, and that the anticipation of regulation also spurs inventive activity."

Firms, being strategic, will innovate in anticipation of new regulations in an attempt to not only be undamaged by the regulation, but also to gain market advantage over their competitors when the regulations are put in place, and thereby increase their profits. In addition, the strictness of the regulation induces concentration along particular technology paths, thereby not only influencing quantity but also type of innovation. Finally, significant impacts occur from government sponsored knowledge sharing activities, such as technical conferences (Taylor, Rubin, and Hounshell 2003).

Case studies have also provided an important method of assessing the innovation impacts of government regulation, and can often better explore the details specific to an individual instance; although lessons learned from one case are not always applicable to others. A consistent result can emerge however, when a large group of cases are in agreement with one another (Becker 2014). For example, a review of a group of 10 environmental regulation studies done from 1970 to 1985 concludes that, 1) "a relatively high degree of [regulatory] stringency appears to be a necessary condition" to induce vigorous innovation, 2) that, "excessive regulatory uncertainty may cause industry inaction," and "too much certainty will stimulate only minimum compliance technology", and 3) when government regulations are signaled in advance, the prospect of their imposition can itself stimulate innovation. Similar findings come from other individual case studies (Ashford, Ayers, and Stone 1985).

Regulation, especially in the environmental industry, acts to secure demand for innovations along particular technology paths, and thereby assure firms of market opportunities, which is a strong incentive to innovate (Mowery and Rosenberg 1982). In this way, regulation can be seen to play a somewhat similar role to positive government inducements to innovate, as it indirectly shifts firms' profit probability distributions; however, wasteful effort is less likely than in the case of subsidies and tax incentives since unsuccessful innovations are not rewarded.

Certainly, not all inventions occur as a result of regulation – in fact, most probably don't. However, regulation can energize a search for new techniques in an industry and thereby spur an innovative burst. This is especially true with regard to environmental-quality related innovations, since there is little else (aside from a desire to reduce waste and materials costs) to motivate less environmentally damaging techniques.

2.2.1.3   Knowledge Stock

For both the world in aggregate and individual firms, cumulative knowledge or "knowledge stock" is a major determinate of innovation, as it directly affects the capacity to generate new knowledge, assimilate external knowledge, or even to understand the value of new knowledge (Cohen and Levinthal 1990). An analogy at the individual level exists, in that a person's level of education influences his or her ability to understand and assimilate new information, and even generate further knowledge.

2.2.1.4   Geography

The importance of industry economic knowledge is the main factor driving the degree of geographical concentration of innovation. Even when controlling for concentration of production, in industries in which skilled labor and knowledge work is more important there is a

greater tendency for innovation to occur in clusters. This is an intuitive result if vital knowledge spillovers are occurring between firms and with universities. Such clustering in industries does not occur in lower levels of knowledge work (Audretsch and Feldman 1996).

Moreover, when looking at patents as a measure of innovation (the use of which I discuss below) diffusion of technology is geographically localized, as indicated by patent citation. "Controlling for other factors [*including language*], within-country citations are more numerous and come more quickly than those that cross country boundaries" (Jaffe and Trajtenberg 1996).

2.2.1.5　　　Organization and Management

The structure and management of organizations doing the innovating plays a role in their creative output. Many studies on individual firms or industry clusters have found a variety of influences of firm structure on innovation. A meta-analysis to examine innovation and its determinants within organizations inspected 13 factors using statistical models and found stable results for, "statistically significant associations for specialization, functional differentiation, professionalism, centralization, managerial attitude toward change, technical knowledge resources, administrative intensity, slack resources, and external and internal communication" (Damanpour 1991).

However, one complicating factor when studying innovation and making policy about innovation, is that it can be highly idiosyncratic, with processes and influential factors varying not only by the field in which the innovation is taking place, but also depending upon the type of innovation that an institution is engaged in. For example, a sustained major factor encouraging technical process innovation within a firm is employment of technical specialists; however, in the 1970 and 1980s, management practices and attitudes did not seem to matter, as they did with social or service related innovations (Dewar and Dutton 1986). More recent research suggests that certain staffing, performance appraisal, and company sponsored knowledge acquisition and sharing activities do have a positive impact on innovative activity (Chen and Huang 2009).

The size of firm can also matter, as larger firms are more likely to suffer communication problems and stagnation in habits that can inhibit innovation (Dougherty 1992).

2.2.1.6　　　Academic Research

Public research is an important determining factor in innovation broadly, since basic science is largely conducted by academic institutions and government labs. In particular industries however, public research is especially important. Pubic research is critical to R&D in the manufacturing and industrial sectors: not only does public research contribute new ideas to industry, but it also "contributes to the completion of existing projects in roughly equal measure overall." (Cohen, Nelson, and Walsh 2002). Moreover, their work explored the channels of communication between industry and the university and discovered that various forms of research results influenced industrial R&D, including papers, conferences, formal meetings, informal channels, and paid consulting. Finally, firm size was important to the degree of impact from public R&D, with large and start-up firms benefiting the most.

Insofar as academic research generates increased innovation activity, factors influencing academic output are important as indirect determinants of innovation. A research project at Louis

Pasteur University using OLS and Tobit regressions found that increasing the size of a lab was a significant negative factor in individual researcher performance, while increasing the amount and quality of other affiliated researchers' output was a significant positive factor affecting individual performance. Moreover, individual characteristics including age and junior status are negatively correlated with research performance, while full-time research status (i.e. no teaching) is positively correlated. The presence of foreign post-docs is correlated with higher publishing, while the presence of domestic post-docs is associated with increased patenting activity (Carayol and Matt 2006). Conversely, in the field of economics, active teaching commitments enhance research quality (Becker and Kennedy 2005).

These results may not be able to be fully generalized; however, many of their covariates provide evidence of the validity of anecdotal observations and research case-studies.

2.2.2 Innovation Outcomes

Various specific and idiosyncratic results are due to individual innovations, but the measures of overall social and organizational impact from innovation are much simpler and more generally applicable. Such outcomes include changes in GDP per capita at the national or international level, and changes in profitability, organizational direction, organizational absorptive capacity, and further innovation cycles at the firm/organization level. Cultural changes are obviously outcomes too, but they are beyond the scope of this work.

2.2.2.1   GDP Per Capita

Increases in productivity of the economy or GDP per capita is the broadest outcome of innovation, and represents an aggregation of all the individual ways that innovation positively impacts humanity. Methods of measuring the contribution of innovation to GDP is addressed in the "measures" section below, but I will address the degree of contribution of innovation to GDP here.

Estimated returns to R&D investment in the G7 and G15 nations' own-rates of return in 1990 were 123 percent and 85 percent respectively, per year (Coe and Helpman 1995). This indicates major and immediate GDP increases from investments in knowledge stock via R&D, often paying back in less than a year. These rates of return are unusually high; however, other scholars have found domestic rates of return in excess of 40 percent, indicating that R&D investment is a highly effective growth strategy.

Moreover, "in 1990 the average worldwide rate of return from investment in R&D in the G7 countries was 155 percent… For the G7 countries the difference between the worldwide and the own rate of return is about 30 percent, which implies a large international R&D spillover; about one quarter of the total benefits of R&D investment in a G7 country accrue to its trade partners." (Ibid.) These spillover gains were stronger in more open economies. "Case study evidence of individual research areas (such as satellites and civilian aircraft) supports the view that social returns to such R&D can be substantial, although extremely difficult to trace and measure... But, again, case studies and the history of individual technologies suggests that these returns are positive and could be substantial." (Hall 1996).

2.2.2.2　　Profitability

The most obvious outcome of innovation at the firm level is an increase in net revenue, either from the creation and marketing of a new product or the improvement (both through cost reductions and quality increases) of an existing one. All of the types of innovation discussed above may result in gains to profit, provided there exists an intellectual property system that will allow a firm to maintain some amount of control over its innovations. Patenting is one such system, which I discuss below.

2.2.2.3　　Organizational Direction

Another firm-level outcome of innovation is augmentation of the knowledge stock of the organization. "What a firm has done before tends to predict what it can do in the future. In this sense, the cumulative knowledge of the firm provides options to expand in new but uncertain markets in the future" (Kogut and Zander 1992). In that way, innovation broadens the set of options of firm behavior and may create new market opportunities for the firm or new organizational motifs.

2.2.2.4　　Absorptive Capacity

Innovation in firms can also have one of the same outputs as public R&D in universities: education. In addition to commercial outputs, innovation in a firm, "develops the firm's ability to identify, assimilate, and exploit knowledge from the environment-what we call a firm's 'learning' or 'absorptive' capacity." (Cohen and Levinthal 1989). Similar to organizational direction, absorptive capacity changes the behavior of the firm, but does so by enabling it to better assimilate information from outside, rather than creating only endogenous changes. Absorptive capacity applies to the full range of outside knowledge, from use of technology products to exploitation of basic scientific and engineering research.

2.2.2.5　　Further Innovation

As befits the autocatalytic nature of research (i.e. answering questions typically generates more questions to be answered), innovation in an organization often leads to further innovation efforts. The phases of implementing new innovations in services yields a "reverse product cycle consist[ing] of a first stage in which the applications of the new technology are designed to increase the efficiency of delivery of existing services: a second stage in which the technology is applied to improving the quality of services: and a third stage in which the technology assists in generating wholly transformed or new services" (Barras 1984). Those wholly transformed or new services constitute innovations in and of themselves.

Not only do technological innovations inspire service innovations, but they open the doors for other technology as well. Without the harnessing of electricity, most modern conveniences would not be able to exist. It is fundamentally those new opportunities that serve as the engine of real per capita growth.

2.2.3 Measures of Innovation

Researchers are fortunate to have multiple measures for the quantity and quality of innovation taking place, including: changes in firm, industry, or overall economic productivity; patenting activity and patent citations; academic publications; intra-firm measures such as percent of the organization engaged in knowledge-work, percent of the budget spent on R&D, and percent of the revenues coming from recent innovations; and, other measures such as direct impacts from a given class of innovations on the environment, and changes in balance of payments between nations.

Some of these measures are rough proxies for innovative activity, such as patents, publications, and the intra-firm effects, while others such as production growth, revenue shift, and various direct impacts come closer to measuring innovation outcome.

2.2.3.1 Production Changes

Both within industries and for economies as a whole, production changes are one of the most direct and easily observed measures of innovative activity. The classic econometric approach involves modeling the impact of the many inputs to GDP (GNP in some papers), with total current knowledge stock (defined as a lagged and depreciating function of prior investments in research) as one of the inputs, and examining how changing knowledge directly impacts GDP (i.e. the partial derivative), and how it impacts GDP both directly and indirectly (i.e. the total derivative – embodying the impacts to other factors as well). However, this approach presents several difficulties including the standard problems with measurement of quality changes and aggregation of dissimilar things (Griliches 1979). Measurement has deteriorated due to the relatively recent complexity of the economy – it used to be agricultural and a growing industrial manufacturing base rather than the highly varied service-driven economy of today – and increasing reticence by firms to answer questions, as well as the rapid rate of change that has made products become noncomparable with ones of the same name from a few years prior (Griliches 1994). Methods of quality adjustment are now employed in an attempt to account for this when calculating GDP growth and price changes; such efforts meet with varied success.

Within firms, endogenous production cost and quality shifts also provide readily observed measures of innovation, and can often be modeled using learning curves. "Quantitative modeling of experience curves has become an increasingly common method of representing endogenous technical change in long-term integrated assessment models used for energy and environmental policy analysis" (Yeh, Rubin, Hounshell, and Taylor 2007). This technique reveals incremental innovation that is difficult to detect using other measures, but suffers from sensitivity to modeling assumptions (ibid.).

2.2.3.2 Patents

Patents present one of the most convenient and transparent measures of innovation. They have received heavy emphasis due to the wealth of data that the patent system collects and the transparency of the information. Moreover, they are a good global measure because international markets generate incentives to harmonize intellectual property laws and patents in industrialized

nations. "Patent data are available for many countries, broken down for the country of origin of the invention for which the patent is requested and for the technological category of the invention. Patent statistics are available in almost all industrialized countries and in many developing countries, broken down for very detailed technological sectors. They probably represent the most precise and reliable classification of technological products… From the name of the body to which the patent is issued, it is possible to reconstruct the sectors of the bodies to which the producers of innovations belong…" (Archibugi 1988).

As with every measure, patents have strengths and weaknesses as an indicator of innovation. Strengths include: provision of a novelty filter, such that they prevent double counting of innovations; provision of a quality filter, such that only inventions that are likely to provide real benefits and therefore be propagated are counted; classification by technology area and sub-area, such that innovation in particular sectors and differential development (both by time and type) is analyzable; citations to prior and later art, such that threads of innovation are detectable and network analysis is possible; and public availability, which increases the transparency of the measure (Archibugi and Pianta 1996).

Weaknesses include: incomplete capture of truly novel inventions, as some are not patentable for various reasons (such as coverage by alternate laws – e.g. copyrights, simplicity, generality, and/or inability to detect infringement) or inventors choose to protect them with secrecy rather than with a patent; differing propensity to patent both by firm and industry creates a degree of non-comparability between patent analyses; differing propensity to patent by nation, and in foreign and domestic markets creates biased distributions when looking within a national patent data set; and idiosyncratic patent laws (despite harmonization efforts) creates a bias when aggregating nations' patent data set (ibid.).

In addition, the value of individual patents is highly skewed, which is a relevant characteristic when there are small counts in an industry or sector (Swann 1993). Value has been assessed in a variety of ways, including: patent citations as indicative of depth and breadth of impact; nations applied in, to measure the geographic influence and generality of an invention; patent claims made, indicating the number and range of novelties; and paid renewal fees, providing a rough indicator of threshold values of particular patents (Archibugi and Pianta 1996).

While firms are most likely to use secrecy and lead time to protect their inventions (Cohen, Nelson, and Walsh 2001), survey work indicates a significant portion of inventions are patented, with 66 to 87 percent of US firms' patentable inventions ultimately submitted (Mansfield 1986). That does not include all of the inventions that do not qualify as patentable, but will nonetheless result in innovation. Therefore, patents represent a good-quality but imperfect proxy for innovative activity in the absence of comprehensive R&D information, and they can be used as a measure for both innovative input and output (depending on the situation). However, given their skewed and highly variable values, patent counts are not a good gauge of R&D output changes in the short run, and should be used cautiously (Griliches 1990).

To complicate matters, the meaning of patents and patent counts has changed over time. The 1970s and early 1980s saw a decrease in US patenting activity despite growth in R&D spending. This was driven by administrative and legal changes which served to tighten the standards of what was patentable, yielding higher average quality of patents (as measured by economic value) (ibid.). Subsequently, starting in 1984, there was a large increase in domestic patenting along with foreign applications in the US patent system, with growth in all sectors, but

concentrated in high-tech industries, where there was a strong increase in profitability of patenting (Hall 2004). This may have been due to the increasingly widespread use of information technology (IT) in two ways: a resulting pressure to protect high-value inventions from imitation or duplication by other firms, and an explosion of high-tech innovation in IT itself. Large firms in particular have shown an increased propensity to patent recently, with patents being used to block patenting by other firms and/or force competitors into negotiations (Cohen et al. 2001).

Other important recent changes to the patent system include: a major influx of foreign applications (Griliches 1994); the advent of a specialized patent court, resulting in improved legal review and higher expected patent values (ibid.); and changing rates of technological development in several sectors (e.g. the previously mentioned IT sector, computers broadly, biotechnology, and pharmaceuticals).

Despite complications and changes, patents remain extremely good "within industry" indicators of firms' relative innovative activity. There are strong statistical correlations between patenting and R&D expenditure both between firms in the same industry and, more weakly, within individual firms over time (Griliches 1990).

2.2.3.3   Publications

Academic publications might serve as a proxy for measuring innovation because innovation relies on results generated by basic research. Similarly, publications, measured by field, can serve as an indicator of the direction of innovation.

The use of publications is supported by their congruence with another measure of innovation: patents. "The probability that the scientist applies for one or more patents during a five-year interval is significantly related to whether or not the scientist has published one or more articles during the five-year period" (Stephan, Gurmu, Sumell, and Black 2002). The number of articles published by a scholar is correlated with their receipt of federal funding, showing a relationship between publications and R&D expenditure (ibid.).

Mixed support exists for publications as a measure of innovation. One paper, "examines the available United States data on academic research and development (R&D) expenditures and the number of papers published and the number of citations to these papers as possible measures of 'output' of this enterprise… for science and engineering as a whole, for five selected major fields, and at the individual university field level," (Adams and Griliches 1996). Using a production function analysis of the data at the field level, they found:

> *…significant diminishing returns to "own" R&D, with the R&D coefficients hovering around 0.5 for estimates with paper numbers as the dependent variable and around 0.6 if total citations are used as the dependent variable. When we substitute scientists and engineers in place of R&D as the right-hand side variables, the coefficient on papers rises from 0.5 to 0.8, and the coefficient on citations rises from 0.6 to 0.9, indicating systematic measurement problems with R&D as the sole input into the production of scientific output. But allowing for individual university field effects drives these numbers down significantly below unity.*

Because in the aggregate both paper numbers and citations are growing as fast or faster than R&D, this finding can be interpreted as leaving a major, yet unmeasured, role for the contribution of spillovers from other fields, other universities, and other countries. (Ibid.)

Although imperfect due to unaccounted-for spillovers, publications seem to represent a useful measure of innovation, especially when used to investigate high-tech fields that may be more closely linked to basic science and engineering work, or industries with strong links to university research. A further challenge to using publications arises in fields where private organizations run large research laboratories (accounting for a sizable fraction of the research activity in the field) and have little incentive to public academic papers. Especially in these areas, within-firm (or intra-firm) measures become more important.

2.2.3.4   Intra-Firm Measures

Four frequently used intra-firm measures of innovation exist. They include, 1) the percentage of workers engaged in innovative activities, or sometimes just the percent of "knowledge-workers" in a firm, 2) the number – often relative to the size of the firm – of innovation-related projects that the firm has completed, is undertaking, and/or is planning, 3) the amount of the firm's budget dedicated to innovation activities, and 4) the percentage of the firm's revenue currently being obtained from products or services innovated within a given number of prior years (OECD 1995).

The first three measures constitute inputs to innovation within a firm, while the fourth is an outcome measure. Other possible outcome measures include: market share gain, profit growth, and/or the return on investment from new products and services.

2.2.3.5   Other Measures

Other measures of innovation tend to be highly individualized to given fields, issues, and/or research questions.

The term "eco-innovation" broadly covers innovations with environmental benefits (Arundel and Kemp 2009). As such, measuring direct impacts on environmental externalities is a legitimate method of measuring innovation in this aggregate field. Moreover, eco-innovations do not necessarily need to be intended to have environmental benefits, but include "all products, processes, or organizational innovations" that do, and therefore measuring environmental benefits not only creates a gauge of innovation but also helps define the field of interest. Specifically, "more effort should be devoted towards direct measurement of eco-innovation outputs using documentary and digital sources to complement the current emphasis on innovation inputs such as R&D or patents." (Ibid.) Also, they state, "Innovation can also be measured indirectly from changes in resource efficiency and productivity." (Ibid.) This is similar to a production-shift method of measuring innovation.

Technology balance of payments (TBP) is a measure of innovation that "measures transactions between firms and sectors of different countries" (Archibugi 1988). In doing so, TBP detects differing rates of innovation, by country and sector, as well as creates an imperfect

measure of technology spillovers. "Among the merits of the technology balance of payments is that of giving data in terms of currency and thus indicating economic relevance of each individual technology transaction." (Ibid.) Weaknesses of the technique include a lack of measurement of input resources to the technology being transferred, an exclusion of non-commercial technologies, and an inability to detect within-firm and within-country technology flows (ibid.)

2.3   Innovation Policy

In this section I will look at the effects of policy instruments on innovation, particularly with respect to environmental innovation because it is an area of importance to the intersection of innovation and policy, and also discuss some complications associated with policies affecting environmental innovation.

2.3.1   Policy Instruments and Innovation: the Pitfalls

Policy tools are regularly employed to incentivize and mold innovation, including: taxes, permits, subsidies, intellectual property laws, public R&D efforts and funding, and direct regulations. The last three will look familiar, as I discussed them in the "Determinants of Innovation" section above; however, here I will go beyond making a case that these mechanisms influence innovation and discuss subtleties and impediments to their fostering the type of innovation that is socially desirable. I will not discuss innovation waivers and technology transfer mechanisms.

2.3.1.1   Market-Based Mechanisms

Taxes, subsidies, and permits are typically referred to as market-based mechanisms and are contrasted with regulation – often favorably by economists – as allowing a more flexible response to mitigating negative externalities or promoting positive ones. Moreover, market-based mechanisms should produce innovations at the lowest cost if calibrated correctly, and thereby produce the optimal net social benefits. However, theoretical results are often ambiguous and depend on assumptions of uncertain validity. "[W]hich policy instruments are most effective in encouraging innovation and diffusion depends upon specific elements of instrument design and/or characteristics of affected firms" (Jaffe and Stavins 1995).

Unless the levels of each market-mechanism are set appropriately, they may not induce innovation at all, or may induce path-dependent innovation trajectories that will not ultimately yield the most socially desirable results. For example, if the marginal social cost is low enough, compared to the profits associated with production and costs of innovation, and the tax reflects the marginal social cost, it will be less costly for firms to pay them and continue production than innovate. This is true even if innovation would be cost effective for the industry as a whole, if there is insufficient coordination to create cooperative research agreements and the R&D costs are too high for individual firms to receive full compensatory benefits. The same result holds even when there is sufficient coordination within an industry, but external linkages are weak, while external spillovers are large. Subsidies suffer from a symmetric problem. Proper

calibration of taxes and subsidies therefore is vital to their efficacy, but as it is often an iterative and opaque process, it is subject to political manipulation and inconsistency. Further, permits (e.g. cap-and-trade systems) have their own problems, including (again) political manipulation by powerful industries, environmental justice issues, and price volatility that can impact incentives to innovate (Taylor 2008a, Taylor 2008b), not to mention complex characteristics such as "offsets" that may or may not be monitorable, verifiable, and/or effective.

Market mechanisms run into further difficulties when issues of path-dependence are considered. Considering the case of pollution: theoretically, if the long-term trends in price and quantity limits associated with pollution are clear, then firms should derive an accurate net-present-value for abatement technology and arrive at an optimal set of investments; however, price trends are highly volatile in practice and are anything but sure when faced with the scientific uncertainty associated with acceptable levels of many pollutants. This is only complicated by political uncertainty and the practical impacts of exogenous events such as global business cycles. Therefore, given the extreme variance of future price estimates, it might be in the interest of firms to pursue short-term "patches" or explore methods to evade regulation (such as outsourcing polluting components of their business abroad), resulting in higher long-term costs to the firm and/or society.

2.3.1.2   Intellectual Property Rights

The design of intellectual property laws can have subtle impacts on incentives to innovate as well. Consider the relationship between "leading" patent breadth, defined as protecting a patent against being supplanted by a new invention, and patent life, defined as the time from application to termination due to expiration or displacement, on diffusion of products and R&D costs: shorter lived but broad patents improve diffusion of novel products, while longer lived and "narrow" patents decrease R&D costs (O'Donoghue, Scotchmer, and Thisse 1998). In addition to the life span and degree of protection provided by intellectual property laws being significant to innovation incentives, the degree of harmonization among different nations' intellectual property laws is also important as it typically serves to strengthen protections overall (Scotchmer 2004a).

A further potential pitfall of intellectual property rights is described by the inversion of a classic economic principle:

> *The "tragedy of the commons" metaphor helps explain why people overuse shared resources. However, the recent proliferation of intellectual property rights in biomedical research suggests a different tragedy, an "anticommons" in which people underuse scarce resources because too many owners can block each other. Privatization of biomedical research must be more carefully deployed to sustain both upstream research and downstream product development. Otherwise, more intellectual property rights may lead paradoxically to fewer useful products for improving human health.* (Heller and Eisenberg 1998)

2.3.1.3   Public R&D Funding

Public R&D funding is another widely used policy tool to spur innovation, particularly

where technologies are expected to mitigate negative externalities. Public R&D is often more politically palatable than prohibitions and effective in generating beneficial and widely available knowledge spillovers. However, analyzing the impact of such funding becomes complex due to those very spillovers, such that the performance of entire industries, rather than impacts to a single firm, becomes relevant (Klette, Moen, and Griliches 2000). Disincentives to utilize public funding become apparent when international spillovers are examined. "There are not institutions to harmonize public spending, and there are no international mechanisms to repatriate the spillovers it generates. As a consequence, there may be too little public sponsorship and too much intellectual property." (Scotchmer 2004a).

2.3.1.4 Regulation

Finally, direct "command-and-control" regulation is as fraught as the other options. "In the presence of weak or nonexistent environmental policies, investments in the development and diffusion of new environmentally beneficial technologies are very likely to be less than would be socially desirable. Positive knowledge and adoption spillovers and information problems can further weaken innovation incentives." (Jaffe, Newell, and Stavins 2004). Parallel problems exist with excessively stringent regulation, as it stifles not only socially desirable levels of production but, perversely, also stifles innovation as the relevant firms simply go out of business or cease producing the impacted product. When regulation takes the form of standards, incomplete information can trap industries into less efficient technology yielding more path-dependencies (Farrell and Saloner 1985). Again, information transparency, predictability, and proper calibration become the issues with regulation, and those face the same impediments noted above.

Another information problem with regulation exists in the asymmetry of knowledge between regulator and regulated firm about the marginal cost of compliance. Firms will tend to make conservative, often exaggerated, estimates of the cost of compliance to try to protect their profits. Several examples of this come from the auto industry where emission controls, seat belts, and fuel economy standards have each been vigorously opposed on the grounds that they would destroy the industry. They did not.

Recent empirical work has highlighted the complexities of using government regulation to foster innovation by pointing out that government actions have both spurred and inhibited the development of three solar energy technologies (Taylor et al. 2007).

Collectively, the difficulties associated with these various policy instruments might inspire despair and complacency; however, policies need not be monolithic and a proper combination of instruments can mitigate the weaknesses of each one individually. Exploring a congruent combination of policy measures is now a major (and highly complex) task for policy scholars.

3   Monte Carlo Simulations Reveal High False Positive Rates with Regression Techniques Typical of Patent Data Research

3.1   Introduction

This chapter will critically examine some of the most common time series methods currently used to analyze patent data, and will demonstrate their higher than expected propensity for generating false positives (i.e. Type I errors). These results will therefore call into question a large number of studies that use such methods and suggest that their conclusions are likely erroneous. Many such studies are used to advocate for policy alternatives, such as increasing or decreasing funding for a particular type of research and development (R&D), or to predict impacts of implementing given policies. Therefore, ensuring that false positive rates are appropriately low when conducting policy-relevant research is itself a policy-relevant issue. Improving the quality and accuracy of policy-relevant research will improve the quality and efficacy of policy making.

Specifically, this chapter will criticize the use of several count-data (non-negative, integer valued) regression techniques that have become popular for use with time-series in the last 20 years, including negative binomial regression (and thereby implicitly Poisson regression) and log-log regressions (which are used to transform positive-valued variables so as to interpret their coefficients as elasticities). Section 3.2.2 will look at several patent-data models using these methods in prominent recent papers, and then discuss why those methods are likely flawed. After that discussion, I will develop a set of techniques for analyzing the reliability of those regression models applied to patent data, and then apply those techniques to regressions with my own patent data (which will be used further in chapter four).

In addition, this chapter will provide an empirical (simulation based) method for excluding forms of analysis from use on time-series[1] patent data. This method may generalize to other types of time-series data, but examining that possibility is beyond the scope of this work.

3.2   Background

This section will provide background details on the differences between stationary and non-stationary time series, known issues with analyzing non-stationary data, and patent data as a type of non-stationary data – or at least data for which non-stationarity is likely and unable to be ruled out.

3.2.1 Non-stationary Data and Regression Modeling

As early as 1974, research indicated problems with using ordinary least squares regressions (OLS) on time series that could be modeled as random walks – specifically, OLS generated high rates of false positives, "spurious regressions," when used on random walks (Granger and Newbold 1974). In a random walk, the prior time period is taken as the starting point for changes in data in the current period, rather than starting from some typical level. This

---

[1] I extend the application of this analysis, in a limited manner, to panel data (a.k.a. Longitudinal, or cross-sectional time-series) in Appendix C.

causes them to "wander" (see Graph 3.2.1.2 below). Equation 1, below, is the classic formulation for the data generating function of a random walk.

$$y_i = y_{i-1} + e \tag{1}$$

In this way, errors persist in their impact over time, accumulating. If the coefficient on the prior period is less than 1 in the data generating function, the impact of the error term from that time period will tend to die out over time; however, if the coefficient equals one, the error's impact will persist permanently. A data generating process which includes the prior period and has a coefficient of 1 on that period is said to have a "unit-root".

Technically speaking, if the characteristic equation of a stochastic process has 1 as one of its roots, that process is said to have a "unit-root," and is also called a "non-stationary" process, since its statistical properties are not constant over time. The variance of a non-stationary time series blows up to infinity as time goes to infinity (Dickey, Bell, and Miller 1986).

"Stationary" processes either do not include the prior period's level in their data generating process (i.e. the data generated in each time period is truly independent of prior periods) or the impact of prior periods decreases with time.

Graphically, the differences between stationary and non-stationary processes are pronounced. Graph 3.2.1.1 (below) shows a randomly generated stationary process with a mean of 1000 and a standard deviation of 200. The line fluctuates and regularly crosses the mean line. Relatively few data points are temporally clustered, and the magnitude in variation of the values seems fairly consistent over time. The line never deviates by more than approximately 2.5 standard deviations above or below the mean, which is reasonable for a 35 data point set.

A consistent data generating process that has no memory of prior level is just as likely to have an "error" above the average level as below in the next period, so it is unlikely to have values clustered above or below the mean for long periods. Without this property, the mean doesn't predict anything about future levels. A memoryless data generating process is also unlikely to produce large deviations from the mean, since each "error" is independently drawn from a normal distribution.

Graph 3.2.1.1: Stationary Series Tend to Fluctuate Around an Average Level
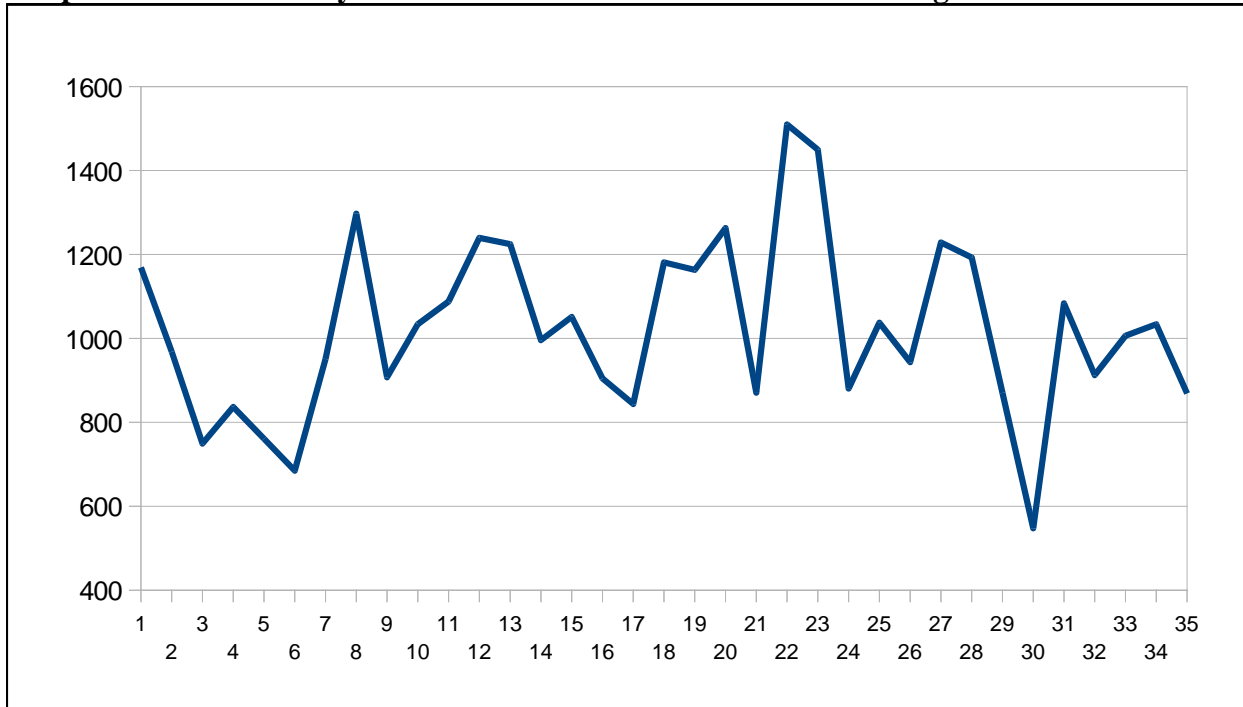
This graph of a stationary random process fluctuates around a mean level of 1000, with a standard deviation of 200. This process shown has a P value of 0.0017 on the Phillips-Perron test, with a trend term, indicating that there is less than a 2 in 1000 chance that the data generating process is actually non-stationary. A OLS regression of this data series against time yields a constant term of 997 (SE: 67), and correctly determines the trend as not being significantly different from zero.

Conversely, graph 3.2.1.2 (below) shows a random-walk: a non-stationary process that I will use later in this essay to generate synthetic patent data for Monte Carlo simulations. The line of the non-stationary graph tends to meander rather than fluctuate, with temporally clustered values and relatively few crossings of the "mean" line. While a mean can technically be constructed for that line, and lies at approximately 77, it has no importance, since the best predictor of the next value of a non-stationary process is the current value (Beck and Katz 2009). While the starting point is 91, subsequent deviations are incrementally small (each error term only has a standard deviation of 12.86) but accumulate to generate large deviations from that constructed mean. The most extreme values of that graph are 120 on the high side, and 21 on the low side, which are 3.34 and 4.35 standard deviations from the mean, respectively. These values are far more extreme than what one would expect in a data set with only 35 points.

Graph 3.2.1.2: Non-Stationary Series Tend to Meander and Have Clustered Values
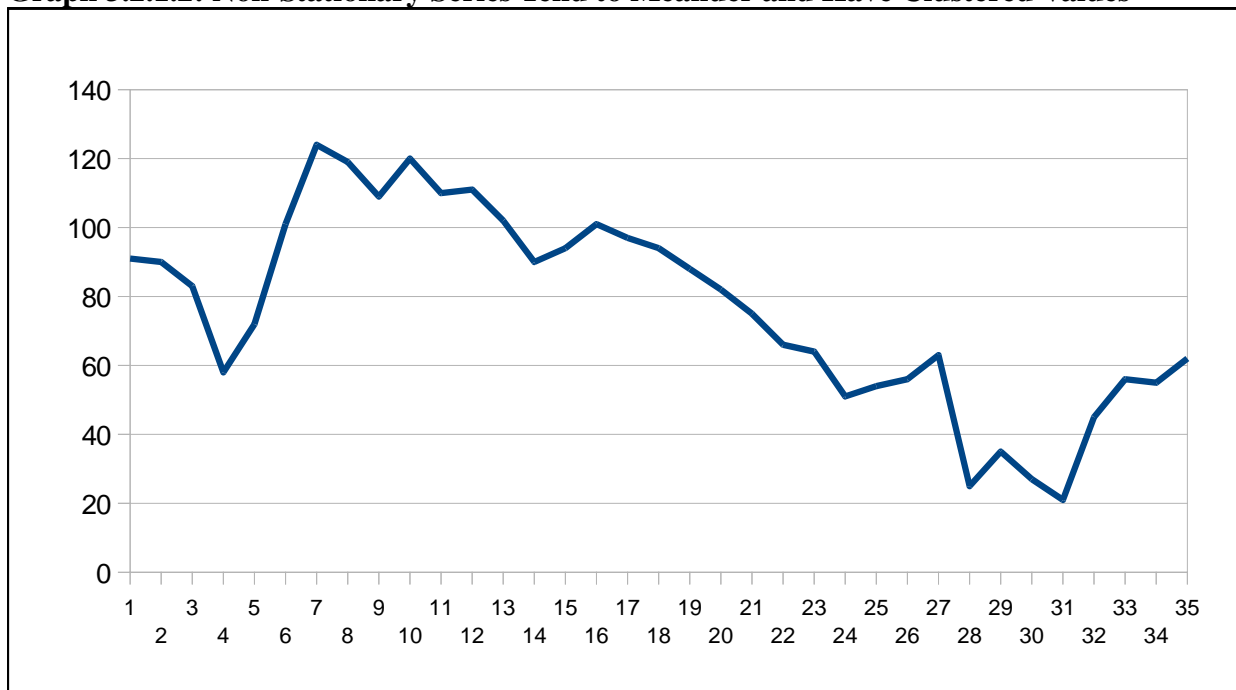
This graph of a non-stationary random process – the same process used in the Monte Carlo simulations below – shows how random-walks tend to meander over time. The first period is a normally distributed random variable with a mean of 60 and standard deviation of 30, yielding an intercept of 91 in this example. Subsequent data points are generated by accumulating error terms distributed as random Gaussian function with mean -0.68 and a standard deviation of 12.86. A Phillips-Perron test, with a trend term, correctly fails to reject the null hypothesis that this data is non-stationary (P=0.55), while an OLS regression of this data versus time produces spurious results: an constant estimate of 113 (SE: 8.5) and a trend coefficient of -2 (SE: 0.38).

       The issue of unit-roots' impact on the statistical validity of simple regressions with "levels" data (i.e. raw measurement data, as opposed to "difference" data where the level of the prior year has been subtracted) was well understood by the mid-1980s (Sargan and Bhargava 1983a and b). Authors used both analytical and simulation techniques to determine that OLS on non-stationary data returned very high rates of false positives and spuriously high R-squared values.

       This problem with using OLS on non-stationary data became a serious concern for econometricians, since the great bulk of macro-economic data could not be definitively shown to not be non-stationary despite decades of scholarship (Nelson and Plosser 1982; Libanio 2005). In fact, while not explicitly non-stationary, many conventional macro-economic models imply non-stationary processes from their results: the Hahn and Solow (1997) model does not yield a fixed or stable "natural rate" of unemployment, and aggregate demand shocks change the long-term path of the model.

       Table 3.2.1.1 (below) shows the results of two unit-root tests on 35 years of observations of three classic macro-economic data series that are often included in statistical analyses as independent variables, including a composite of energy prices, US Gross Domestic Product (USGDP), and the price level of the S&P 500 composite stock index. Both tests strongly fail to reject the non-stationarity null hypothesis on all three series. However, in one sense, perfect non-

stationarity cannot be strictly true for macro-economic data: we would never expect any of these series to actually meander astronomically high as time went toward infinity, nor would we expect their variances to grow to infinity (Beck and Katz 2009). Moreover, tests of non-stationarity have great difficulty distinguishing between actual and nearly non-stationary processes in a finite data set (West 1988). These tests can also be sensitive to the number of lags included, the inclusion or lack of intercept and trend terms[2], and/or the existence of structural changes (Smith 1999).

Table 3.2.1.1: Macro-Economic Data Often Appear Non-Stationary

	Augmented Dickey Fuller	Phillips-Perron
Energy Prices*	0.99	0.99
USGDP	0.71	0.70
S&P 500	0.70	0.51

The two most commonly used tests for non-stationarity – the Augmented Dickey Fuller test and the Phillips-Perron test – fail to reject the null hypothesis of non-stationarity on frequently used macro-economic data, including 35 years of energy prices (2000 dollars), USGDP (2005 dollars) and stock prices (S&P 500 Index price in 2008 dollars). *Energy prices here are a consumption weighted basket of real prices for fossil-fuels used in electricity production.

Nevertheless, a short-run model that accurately approximates the statistical behavior of these series may have unit-roots in its construction. A model that may be false in the asymptotic case may still be locally useful (e.g. Newtonian physics). The behavior of the energy price composite shown in Graph 3.2.1.3 (below) is remarkably similar to the graph of the non-stationary synthetic data above. As with the non-stationary synthetic data, the standard deviation of the changes in energy prices (220) is far smaller than the standard deviation of the levels data (615), the values are temporally clustered, the series shows a meandering pattern, and – using the standard deviation of the changes – the extreme values fall very far from the mean value of $1,997. Therefore, concerns about the impact of non-stationarity on statistical validity are applicable to data series that show similar traits in the short run, even if the variance of the series will not actually grow to infinity.

---

2   Merely removing trends from data by controlling for time is often insufficient to correct for non-stationarity, and the imposition of a trend in "difference stationary" data can even generate spurious results itself (Raffalovich 1994).

Graph 3.2.1.3: Energy Prices Exhibit Typical Non-Stationary Characteristics
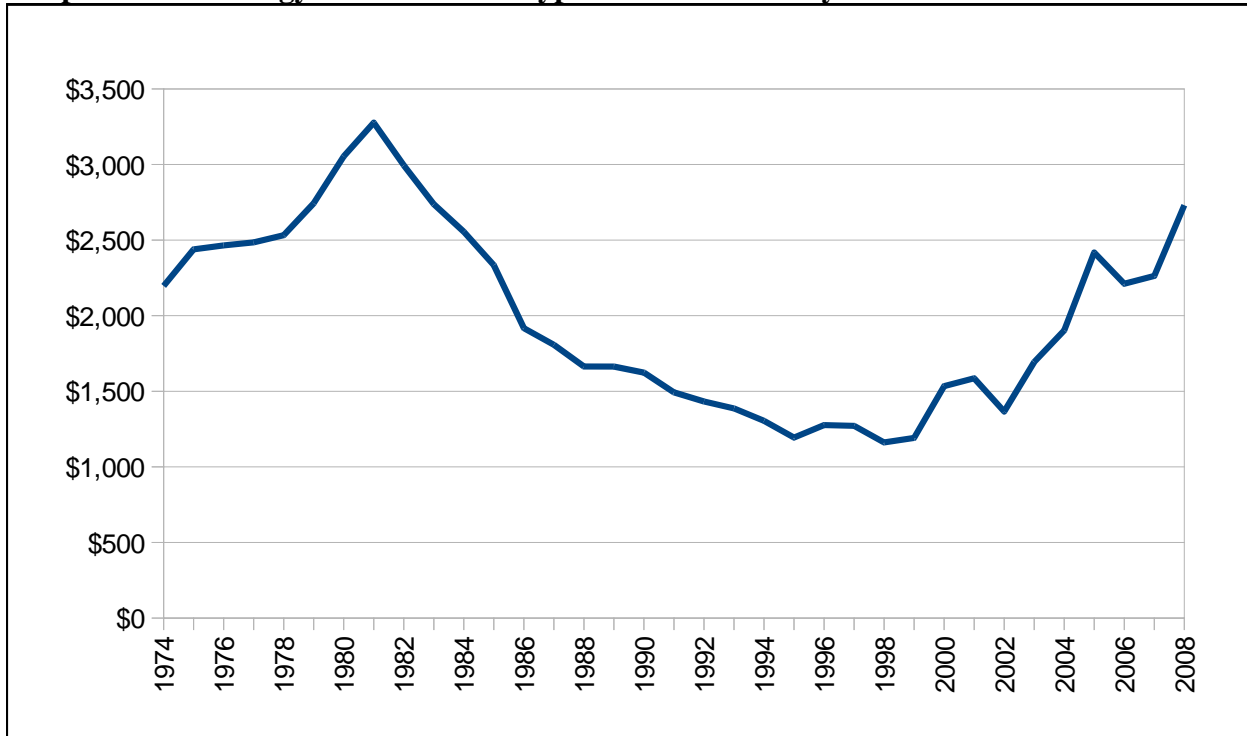
A consumption-weighted basket of prices of fossil-fuels used in electricity production meanders in a manner typical of non-stationary data. Moreover, there is no reason to assume such data would be stationary since energy extraction technology, mix of fuels, and scarcity of resources changes with time, yielding no stable long term average.

Despite the well understood impact of non-stationary data on the validity of OLS results, little to no work has been done to explore similar issues with other regression equations such as Poisson or negative binomial. Unlike OLS, these models do not assume normality of the dependent variable and have therefore been useful for statistical exploration of count-data, which violate the normality assumption (Brandt, Williams, and Fordham 1998). Patent data is a prominent example of count-data.

3.2.2 Patent Data and Other Count-Data Models

Innovation scholars who use patent data regularly employ time-series or panel data in their analyses. The value of using such data is its capacity to control for within-firm effects – such as innovative productivity (Lerner and Seru 2015), reveal the impact of policy events or other treatment effects, and unique ability to chart the evolution of an industry's intellectual property or progress of a sector through the stages of some radical technological change (Katila 2000).

Early researchers using regressions to analyze patent data were frustrated by the fact that count-data (non-negative, integer valued data) provided statistically unreliable results when using OLS (King 1988, 1989a, 1989b). From this discovery emerged an initial preference for using Poisson models with patent counts over time (Brandt et al. 1998), and eventually negative binomial regressions due to their lack of constraint by the assumption of an equal mean and

variance (Cameron and Trivedi 1990, Guo 1996). Also, in the case of mean and variance equality, negative binomial regressions return identical results to Poisson regressions (ibid.).

As a result, negative binomial regressions and related forms (such as zero-inflated negative binomial regressions) grew in frequency of use and popularity through the late 1990s and early 2000s. Macroeconomic and organizational-behavior researchers found negative binomial models particularly useful, since many of the data they used (such as GDP, capital investment, R&D investment, employment, patents, oil prices, etc.) were count-data (or could easily be treated as counts). These data were analyzable with the negative binomial form, and that form helped avoid OLS which had been so effectively criticized.

Many key and highly cited papers from this period (with hundreds, often thousands of citations by other research), including ones on innovation clusters (Baptista and Swann 1998), patent-data models (Blundell, Griffith, and Van Reenen 1995), industry process studies (Benner and Tushman 2002), inter-firm alliances (Stuart 2000), and environmental innovation (Brunnermeier and Cohen 2003) used negative binomial regression models to find new and highly significant correlations.

Many, more recent, prominent patent-data research papers (Lee 2006, Johnstone et al. 2010, and Popp 2005) rely upon negative binomial models, or results from other papers using negative binomial models, to determine patenting-related relationships. These modeling efforts have been generally well regarded: Jaegul Lee won the 2006 Alfred P. Sloan Foundation Dissertation Award for his study of automotive emissions control technologies, which utilized negative binomial regressions on automotive patent data.

Other alternative specifications have been employed, such as log-log regressions, to avoid the problems with over-dispersion, and to allow for interpretation of coefficients as elasticities (Popp 2002). Log-log regressions use OLS, but take logs of both independent and dependent variables before doing the regression. This means that changes in the independent variables are interpreted as percentages, and their impacts on the dependent variables are also taken to be percentages: hence the elasticity interpretation.

Importantly, much of the current patent related research uses panel or time-series data to establish relationships with other series such as prices, research funding, political variables, and/or regulatory events (Lee 2006, Blundell et al. 1995, Johnstone et al. 2010, Popp 2005, and Popp 2002). When using macro-economic data such as these, concerns about non-stationarity raised in the previous section become relevant again.

While working with data on patenting rates in renewable energy technologies, I first discovered that all of the data series I had compiled were non-stationary (see Graph 3.2.2.1 and Table 3.2.2.1, below), and therefore were likely yielding spurious results in my OLS regressions. As I read other patent-based papers, I encountered the regular use of negative binomial models and log-log models in preference over OLS due to their aforementioned relevance to count-data. My subsequent use of negative binomial models returned similarly high R-squared values and similarly significant coefficients to my earlier work, and I became suspicious. Taking first differences of the data resolved the non-stationarity problem but also dramatically reduced the number of significant results my models generated.

In retrospect, discovering non-stationarity in the data was unsurprising: there is no logical reason to believe that there is a "normal" level of patenting in a given technology that should be returned to over time; no mean level of patenting to revert to. A given technology may be in any

number of states, from early stages of discovery – aka breakthroughs, to initial development and popularization (where product innovations dominate), to later development and evolution into maturity (where process and application innovations dominate), to eventual decline and obsolescence (where few if any innovations occur) (Abernathy and Utterback 1978, Anderson and Tushman 1990, Tushman and Rosenkopf 1992). Insofar as patenting reflects innovative activity, patenting may similarly rise and fall based on the technology life-cycle. Given the eventual obsolescence of many or most technologies, the "long run" patenting rate for any particular one might be zero. Sectors also rise and fall with time (e.g. computers and sailing ships) even when their patenting rates never fall to zero; therefore, establishing a "mean" level of patenting for some technology or sector might not be possible.

Graph 3.2.2.1: Patent-Data Series for Four Renewable Energy Technologies
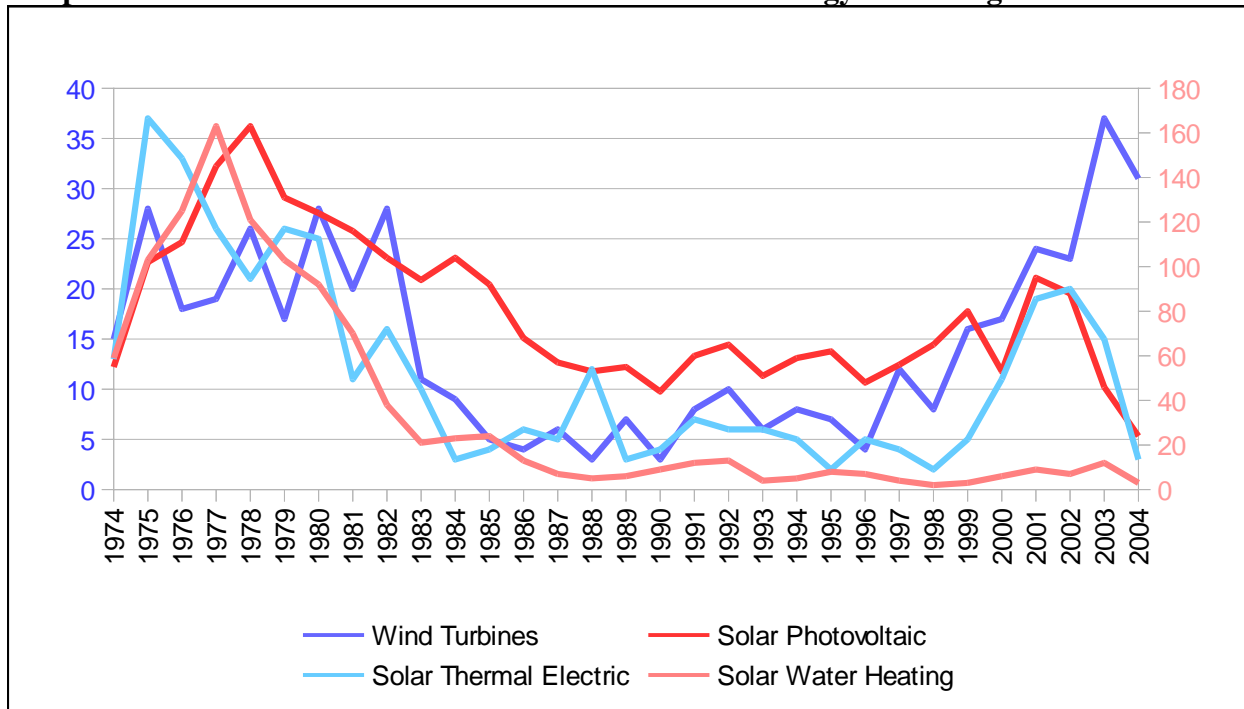
Table 3.2.2.1: All Renewable Energy Technology Patent Series are Non-Stationary

	Augmented Dickey Fuller	Phillips Perron
Wind Turbines	0.76	0.82
Solar Photovoltaic	0.10	0.07
Solar Thermal Electric	0.31	0.21
Solar Water Heating	0.76	0.50

Augmented Dickey Fuller and Phillips-Perron tests, with a trend term, fail to reject the null hypothesis of non-stationarity for any of four renewable-energy technology patent series, although they come close in the case of Solar Photovoltaics.

25

In fact, patenting is deliberately not a memoryless process. Patents specifically refer to "prior art" in a technology field (or related field) in order to appropriately narrow their application (lest they be determined to be overly broad), provide context, and support their claims to novelty. Patent data can be thought of as *intentionally* non-stationary. The only remaining question then is if it can be appropriately thought of as a random walk.

Below, I'll describe my methods for generating random walks that mimic the statistical characteristics of my renewable-energy patent-data. I will then detail the results of using those synthetic data series in Monte Carlo simulations, and show how synthetic patent-data produces high rates of spurious regressions.

3.3   Hypotheses

Null Hypothesis: Non-stationary count-data, and other data with statistically similar characteristics to patent series, are not producing higher than appropriate rates of spurious correlations with their time-series covariates when used in popular regression equations, including negative binomial and log-log.

Alternative Hypothesis: Non-stationary count-data, with similar statistical characteristics to patent series, produce a higher than appropriate rate of spurious correlations when used in negative binomial and/or log-log regressions – as they are know to produce when used in OLS regressions.

3.4   Methods

The central question of this essay is: Do regression techniques typically applied to non-stationary time-series patent data generate spurious results similar to those OLS is known to produce? Since Monte Carlo techniques have been successfully used in the past to test such issues (Granger and Newbold 1974), I decided to generate random walks as synthetic "patent-data" (with similar statistical properties to the actual renewable energy patent-data I was using for the second essay of my dissertation) and test them using a variety of regression models against non-stationary macro-economic data (specifically, United States Gross Domestic Product [USGDP] – a frequently used co-variate in regression analysis – and a non-stationary price series I had generated for the renewable energy patent-data research – a consumption weighted basket of fossil-fuel prices as represented in electric power generation). If the synthetic patent-data based regressions produced more than the expected rate of false-positives (5 percent when using a standard 95 percent confidence interval for the coefficients of the covariates), then I would be able to reject my null hypothesis, and confirm that patent-data regressions using non-stationary covariates were vulnerable to higher than expected spurious correlation rates. I also generate stationary synthetic data as a control.

3.4.1 Non-Stationary Synthetic Data Construction

The general form of the synthetic data used in the simulations is a random walk, described by the equations:

$$y_1 = a, \text{ and for all } t>1, y_t = y_{t-1} + e \text{ or } 1 \text{ if } y_t < 1 \qquad (2)$$

The use of a non-zero centered first period "a" helps both align the synthetic data with the statistical properties of the renewable-energy patent-data and ensure that few of the simulations have synthetic "patent" series that immediately fall to zero and stay there. The "a" term is a normally distributed random variable with a mean of 60 and a standard deviation of 30. Using this random intercept generates synthetic series that span the range of intercepts of the renewable-energy patent-data the synthetic data is modeled on.

After the first period, the random-walk component of the synthetic data begins, with each period accumulating the error terms from prior periods. The error terms are normally distributed random variables with a mean of -0.68 and a standard deviation of 12.86. These values were selected to generate a set of synthetic series such that the statistical properties of the set of these series would span (within two standard deviations) the statistical properties of the four renewable-energy patent-data series. The statistical properties of both the synthetic series set and the renewable-energy data are presented in Table 2.5.1 in the next section.

Rounding up negative or zero values to 1 serves three purposes: first, it reflects the fact that it is not possible to have a negative patenting rate (count data is zero truncated); second, it generates similar patterns to the renewable-energy patent-data the synthetic data is designed to mimic (patenting in several of those series falls to low levels – 2 patents per year – but never fully to zero); and third, rounding up to one allows log transformation of the synthetic data for the log-log regression Monte Carlo simulations (the log of zero is undefined). Larger aggregations of patent data (the entire set of US patents, a given patent class, or even a sub-class) never fall to zero, but most patent-data research is conducted on narrowly defined technologies or technology areas that may experience large drops in patenting, especially in response to policy events or growing patenting in a competing technology.

Finally, rounding values to the nearest integer reflects the fact that patent data are count values, i.e. you cannot have a fraction of a patent.

3.4.2 Testing the Non-Stationary Synthetic Data and Comparing it to Real Data

I compute the average mean, average standard deviation, average minimums and average maximums of the set of 10,000 synthetic patent-data series, with standard deviations of each average statistic. These test statistics should be relatively close to the values of the renewable-energy patent-data series, i.e. within two standard deviations.

Moreover, comparison between the results of various regression types should be similar for the non-stationary synthetic data and actual renewable-energy data. In other words, if OLS, negative binomial, and log-log regressions produce similar results for the actual data series, they

should also produce similar results for the synthetic data; if they produce markedly different results for the actual data, they should do likewise for the synthetic data. If this is not the case, it is likely that the synthetic data does not accurately mimic the behavior of actual patent data.

3.4.3 Generating Stationary Synthetic Data as a Control

In addition to the non-stationary synthetic data, I generate a stationary synthetic data set, with each period being a normally distributed random variable with mean of 1000 and standard deviation of 200 (so it will be unlikely to fall below zero). I then conduct Monte Carlo simulations regressing the stationary synthetic data against macro-economic data. These regressions should produce the expected rate of false-positives (approximately 5 percent), and should produce similar rates of false-positives to the first-differences of the non-stationary synthetic data. Failure to achieve these results will suggest problems with the synthetic data or other elements of my approach.

3.4.4 Selecting Regressions to Use with Synthetic Data

I include OLS – known to generate spurious results – as a test to show that high rates of false positives are indeed possible with these series. If false positives are close to the expected level of 5 percent, with OLS on the levels data, then a flaw in the construction of the series would be apparent. Moreover, I include OLS on the first differences of the data as an inverse of the previously mentioned test: the first differences of 98.5 percent of the synthetic data test as stationary, so high rates of false positives should not occur using first differences, and if they do it points to a flaw in the synthetic data construction.

Due to their previously detailed popularity in patent-data models, I also include negative binomial and log-log regression equations.[3] If false-positive rates are similar to elevated false-positive rates for the OLS models, problems with using negative binomial and log-log specifications with patent-data will become apparent. Conversely, if false-positive rates are close to 5 percent, and similar to levels seen with stationary synthetic data, then the appropriateness of those model specifications will be revealed. While I would like to also analyze first differences using negative binomial and log-log models, the presence of negative numbers in the first differences precludes use of the negative binomial distribution, and the presence of zeros precludes the taking of logs.

3.5 Results

3.5.1 Comparing Synthetic Data and Renewable-Energy Patents

The statistical characteristics of the non-stationary synthetic data sets compare well to the

---

3 Monte Carlo simulations for the negative binomial regressions required segmentation and use of multiple seeds for STATA's random number generator. Large numbers of simulated regressions (typically upwards of 500) tended to eventually generate one regression that would become "trapped" in a convexity or discontinuity during successive iterations and thereby foul the entire run. Segmentation allowed for smaller runs that avoided fouling, from which I then summed the results for a full total of 10,000 runs.

characteristics of the actual renewable-energy patent data: i.e. the actual means, standard deviations, minimums, and maximums all fall within a two standard deviation range of the synthetic data statistics. Moreover, the coefficients of variation (CV) for the actual data span the range from 0.42 to 1.16, with a mean value of 0.76. The mean CV for the synthetic data (0.83) falls close to the mean of the actual data CVs, and the standard deviation of the synthetic data CVs (also 0.83) indicates that CVs for all four actual data series fall comfortably within the range of simulated data series. See Table 3.5.1.1 below.

Table 3.5.1.1: Statistical Characteristics of Energy Patent Data and Simulated Data

	Mean	Standard Deviation	Minimum	Maximum	Coefficient of Variation
Wind Turbines	14.77	9.53	3	37	0.65
Solar Photovoltaic	79.68	33.26	24	163	0.42
Solar Thermal Electric	11.77	9.69	2	37	0.82
Solar Water Heating	34.74	40.33	2	163	1.16
Simulated Data*	55.66	23.36	20.13	101.78	0.83
(Standard Deviations)**	(41.9)	(10.43)	(28.24)	(49.52)	(0.83)

*Each value for the simulated data is the average of 10,000 simulated data sets, i.e. the "mean" reported here is the average of all 10,000 means, the "standard deviation" is the average of 10,000 standard deviations, etc.
**Standard deviations shown here are for the averages of the simulated data above them.

      The next point of comparison between the actual renewable-energy data and the synthetic data concerns the results of the OLS, negative binomial, and log-log regressions. Table 3.5.1.2 (below) shows that each regression type generates a similar pattern of significant (or not) results when the actual data is regressed against USGDP (2005 dollars) and my consumption-weighted basket of fossil-fuel prices (2000 dollars). For example, when wind turbine patenting is regressed against USGDP it fails to generate significant coefficients with any of the three techniques. All other regressions generate significant results. Moreover, the p-values for each regression are of a similar magnitude. The non-stationary synthetic data portion of this comparison will be demonstrated in the next section as I present the results of the synthetic data Monte Carlo simulations.

Table 3.5.1.2: Regressions for Energy Patent Data and Macro-Economic Variables

	OLS		Negative Binomial		Log-Log	
	USGDP	Energy Prices	USGDP	Energy Prices	USGDP	Energy Prices
	P-Value	P-Vale	P-Value	P-Vale	P-Value	P-Vale
Wind	0.690	0.002	0.678	0.001	0.725	0.002
PV	0.001	0.000	0.000	0.000	0.002	0.000
STE	0.029	0.002	0.026	0.000	0.013	0.000
SWH	0.000	0.000	0.000	0.000	0.000	0.000

3.5.2 Monte Carlo Simulations using Synthetic Data

Table 3.5.2.1 shows the regression results for non-stationary synthetic data Monte Carlo simulations. In 10,000 simulations, negative binomial, log-log, and OLS all show very high false-positive rates when conducted on non-stationary synthetic dependent variables with non-stationary actual independent variables. False-positive rates range from 59.7 percent to 78.9 percent, depending on the technique and covariate. This range is consistent with data from Granger and Newbold (1974) for OLS. Moreover, the percentages of false positives for each independent variable are similar across all three regression types: this completes the comparison between actual renewable-energy data and the non-stationary synthetic data begun above – regression results are consistent between techniques on the non-stationary synthetic data, just as they were consistent between techniques with the actual data.

Negative binomial, log-log, and OLS regressions on the stationary synthetic dependent variable show false-positive rates nearly in line with expectations in the case of both independent variables (see Table 3.5.2.1 below). False-positive rates range from 5.21 percent to 5.83 percent, which are slightly elevated over the expected rate of 5 percent, but not egregiously so, and are likely due to forcing a lower bound on the synthetic data. Similarly, the first-differences (labeled "Changes" in the table) on the non-stationary synthetic data – which creates stationary results – show false-positive rates of 5.61 and 5.67 percent for the two covariates in the OLS regressions. These are similar to the rates for the stationary synthetic data.

Table 3.5.2.1: False-Positive* Rates for Synthetic Data and Macro-Economic Variables

	OLS		Neg. Binomial		Log-Log	
	USGDP	Energy Prices	USGDP	Energy Prices	USGDP	Energy Prices
Non-Stationary Synthetic Data	78.90%	61.80%	78.84%	59.70%	78.39%	59.73%
Stationary Synthetic Data	5.71%	5.39%	5.83%	5.64%	5.55%	5.21%
Changes on Non-Stat. Data	5.61%	5.67%				

*False positives here are P values less than 0.05.

3.6   Conclusions:

Negative binomial and log-log regressions have become popular with scholars who analyze patent data, but the validity of their research is seriously called into question by the results presented above. When synthetic patent data is regressed in a Monte Carlo simulation against common covariates like USGDP and a price series (themselves non-stationary), exceptionally high false-positive rates are observed. This indicates a sizable fraction of the policy-relevant patent and innovation literature may rest on very shaky ground.

The results also demonstrate the capacity for synthetic random walks to accurately mimic the behavior of patent-data series. While some researchers have questioned the ability to detect true non-stationarity and/or the importance of non-stationarity in macro-economic data (West 1988, Smith 1999), my results show that it isn't necessary to prove that patent-data series (or any other data for that matter) are random walks, it is sufficient to show that they cannot be distinguished from random walks in order to call into question the results of papers that rely on these techniques. With false positives likely in the range of 60 to 80 percent, the inability to rule out a random walk as the data generating process means we cannot trust results from these regression methods when used with non-stationary covariates on patent data. It also may explain the popularity of some of these techniques, since using them has a high probability of generating spurious "statistically significant" results.

Finally, these results demonstrate the potential for Monte Carlo simulations to be used as a validating method when concerns about the legitimacy of a more exotic statistical technique arise.[4] The question can be explored by generating a large number of synthetic data sets with similar statistical properties to the dependent variables and regressing them against relevant independent variables. When false positive rates are found to be appropriate, researchers can proceed with more confidence.

---

4   Some researchers claim that panel data (i.e. longitudinal) is the solution to the problems I present here. In the asymptote (i.e. as t [the time index] and i [the identification index] go to infinity) non-stationarity stops being a problem (Phillips and Moon 2007); however, given the typical size of panels used in statistical research, non-stationarity continues to be an issue. See Appendix C for a demonstration of this.

4   Public R&D Investment and Renewable Energy Invention

4.1   Introduction

Several highly regarded and highly cited papers have established a strong relationship between research and development (R&D) expenditures and invention (as measured by patenting), both between firms at a given point in time, and longitudinally within a firm (Hall, Griliches, and Hausman 1986; and Griliches 1990). Moreover, these relationships were documented to be highly explanatory, with R-squareds of 0.9 in the cross-sectional case and 0.3 in the longitudinal case, and fairly contemporaneous with very small lag effects if any (ibid.). However, these results were for private R&D expenditures within the manufacturing sector, prompting the question: do these results extend to all sectors, technologies, and/or funding sources?

Margolis and Kammen (1999), claim that public R&D funding for renewable energy technologies and patenting in those sectors have been correlated, and that reduced public funding in those technologies is likely to inhibit innovation therein. This sentiment is echoed in Nemet and Kammen (2007). Popp, Newel, and Jaffe (2009) go further and claim that:

> *As environmental economic research on technological change has grown, the*
> *importance of considering market failures for knowledge, as well as traditional*
> *environmental externalities, has been emphasized. In particular, calls for*
> *increased government support for environmentally friendly R&D are motivated*
> *by the need to overcome such market failures.*

While these statements seem logical on their surface, there is no reason we should presume they are correct without supporting evidence. Also, while casual inspection of R&D funding for renewables graphed with patenting levels suggests a correlation, a slightly more detailed examination reveals that annual patent counts sometimes rise ahead of increases in funding. This is further complicated by the fact that total funding amounts are recorded in the first year of a multi-year grant since the money has been allocated by the Federal government even if it has not been dispersed, meaning that much or most of the money may not be distributed for several years after the funds are recorded as having been given (USDOE 1980). This compounds the observation that patenting levels may rise ahead of funding increases: those funding increases may not actually be fully experienced by the recipient organizations for several years after patenting levels are observed to rise. An empirical investigation of the relationship between federal R&D money and patenting in renewable-energy technology is therefore warranted.

Within each of these papers, as well as within countless popular articles discussing renewable energy and the role for government, there are explicit or implicit causal links claimed between government R&D spending and rates of invention within a given technology. A glace at graphs showing patenting rates and R&D expenditures seems to validate those links.[5] Nevertheless, these claims deserve further examination, especially given the growing importance of developing alternative energy sources in fighting climate change, oil dependency, and the

---

5   See the first chart in Nemet and Kammen (2007).

effects of resource scarcity.

While Hall, Griliches, and Hausman (1986) found that patenting success was not predictive of future R&D spending levels within a given firm, the relationship between public R&D funding and patenting is likely to be substantively different. In the within-firm case, decisions about patenting are primarily driven by expected profits from having exclusive rights to inventions (or protecting a stream of profits from competing intellectual property claims), and therefore one would not necessarily expect past inventiveness to affect future research spending. Conversely, patenting serves a broader set of purposes for public entities or public-private partnerships. Patents may serve as evidence of research success, help career advancement, inform grant review decisions, and form the basis for future grant applications and/or program development. In other words, patents may have a more complex relationship with R&D funding where public money is involved, and patents might therefore drive future R&D levels.

4.2   Background

In this section, I explore the origin of public interest in renewable-energy technologies, the growth and decline of federal R&D funding, and general development of four renewable energy technologies: wind turbines, solar photovoltaic cells, solar thermal electric systems, and solar water heating systems.

I then describe the results of four (lightly) structured interviews with experts in the area of renewable-energy technology and technology funding. I address issues of R&D investment decision making in renewable-energy technology firms, impressions of the role of intellectual property types, and specific motivations to generate patents.

4.2.1 Wind Power

The first federal research efforts for wind power focused on funding aerospace corporations to develop high-output horizontal axis wind turbines, building on existing airplane technologies (Sawin 2001). This approach originated with the Federal Wind Energy Program (FWEP), passed in 1973 in response to the oil crisis. In the decade prior to 1973, federal R&D investment in energy had been overwhelmingly directed at nuclear power – specifically the further development of fission reactors (Dooley 2008) – and renewables were a relatively low federal priority.

The FWEP was primarily concerned with developing commercial-grade wind turbines as fast as possible (Raytheon 1980), and to do so it aimed to promote innovation, bolster wind turbine manufacturing, and diminish other barriers such as negative public perceptions and environmental externalities (CEC 1981). In 1977 the United States Department of Energy (USDOE) took over administration of all wind-power funding and development programs (USDOE 1980).

In 1980, Congress passed the Wind Energy Systems Act, which set three primary goals to be achieved by the end of 1988: 1) to advance technologies that would make wind-power commercially competitive and widely available, 2) to lower the price of wind generated electricity to the level of conventionally generated electricity, and 3) to install at least 800 megawatts of wind generated electrical capacity. The early years of funding were focused on

technology development and wind resource evaluation, with the build-out to occur toward the end of the program (WESA 1980). However, much of the money allocated was never actually distributed, due to the Reagan Administration cutting renewable-energy funding. By 1986, turbine failures and maintenance costs drove research efforts toward increasing reliability, longer component lifespan, and higher cost-effectiveness (SERI 1987).

During the early years, from 1974 to 1981, annual federal funding increased each year, peaking at $160 million per year; however, following the Reagan cuts, funding rapidly fell to under $65 million per year, where it stayed through 2008. Federal funds dipped as low as $17 million per year during that period and never rose above $62 million. See Graph 4.2.1.1, below.

Graph 4.2.1.1: Wind Turbine Patenting Rates and Federal R&D Spending
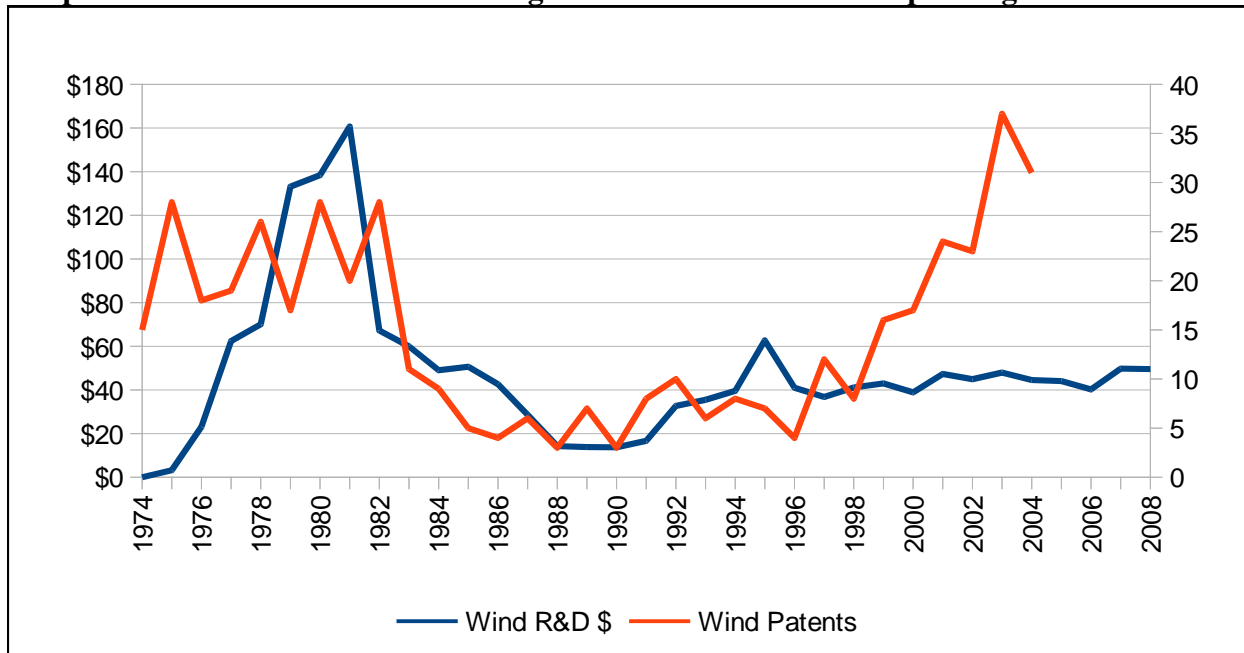
4.2.2 Solar Power

In 1974, congress expanded the scope of the FWEP to include solar power. The initial expansion envisioned an aggressive program that would invest one billion dollars (4.9 billion 2016 dollars) in new technologies (SERDDA 1974). That same year, the Energy Research and Development Administration (ERDA) was created by the Energy Reorganization Act of 1974, and solar technologies were explicitly made a main focus of the new agency (ERA 1974). In 1977, ERDA solar projects and funding was folded into the USDOE as the Solar Energy Research Institute, which later changed its name to the National Renewable Energy Laboratory (NREL) (USDOE 1980).

 Solar power consists of three dissimilar technologies: 1) solar photovoltaic cells (PV) that convert sunlight directly into electricity via photons acting to excite electrons in a material, and thereby induce a current; 2) solar thermal electricity (STE), which consists of reflective panels or troughs that concentrate sunlight on an vessel containing a liquid or gaseous material

(often molten salts) that will then use that heat to power a turbine; and, 3) solar water heating (SWH), the simplest of the three technologies, which typically consists of a tank or piping system, painted black, and mounted on a roof so as to heat water (or pre-heat water that will then flow into a conventional water heater, and thereby reduce energy required to fully heat it [Block and Harrison 1997]) as it is circulated. More detailed descriptions of the three solar technologies and their mechanisms of operation can be found in Taylor et al. (2007).

At the early stage of federal involvement, solar PV and solar thermal electric generation was far more expensive than wind or solar water heating, with costs frequently two orders of magnitude (or more) higher than those of coal fired generation. "The OTA [Office of Technology Assessment] analysis also suggested that the federal effort to develop photovoltaic technologies was misplaced given the nation's urgent need for increased domestic energy production and recommended that the solar program focus instead on solar water and space heating which were technologically and economically "available now" (OTA 1978)." (Dooley 2008).

Despite the differences in readiness for commercialization, all three technologies received approximately equivalent levels of funding, and saw nearly identical patterns of funding growth and decline during the early years: allocated dollars rose sharply and monotonically throughout the mid to late 1970s and peaked at approximately $350 million in either 1979 or 1980 (depending on the technology). Funding for each solar technology declined precipitously thereafter. Renewed interest in solar technologies was accompanied by a small resurgence in funding for PV and STE in the mid-1990s during the Clinton administration. See Graphs 4.2.2.1, 4.2.2.2, and 4.2.2.3 below.

Graph 4.2.2.1: Solar PV Patenting Rates and Federal R&D Spending
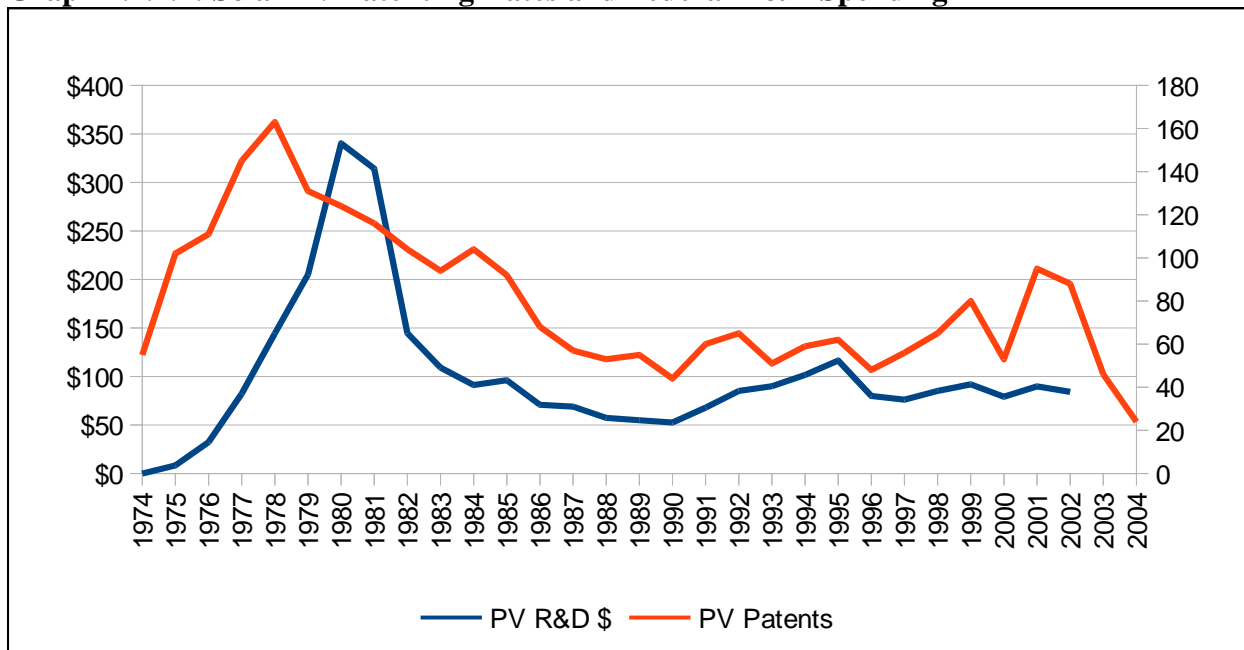
Graph 4.2.2.2: Solar Thermal Electric Patenting Rates and Federal R&D Spending
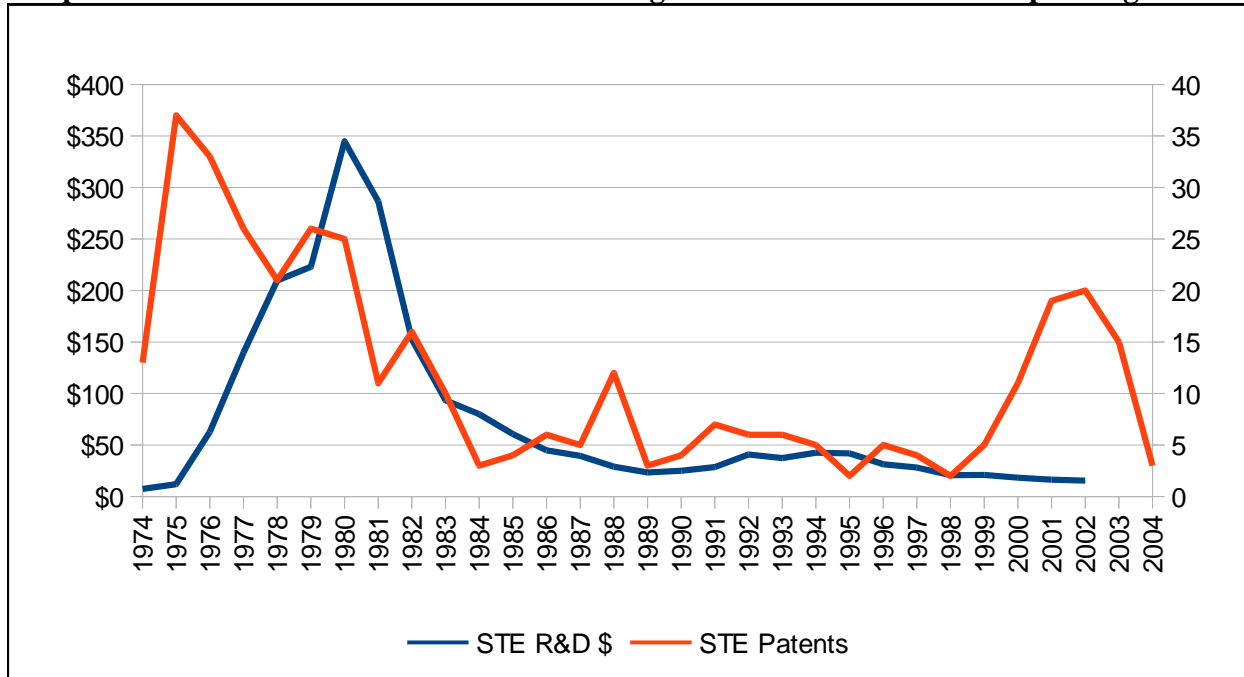
Graph 4.2.2.3: Solar Water Heating Patenting Rates and Federal R&D Spending
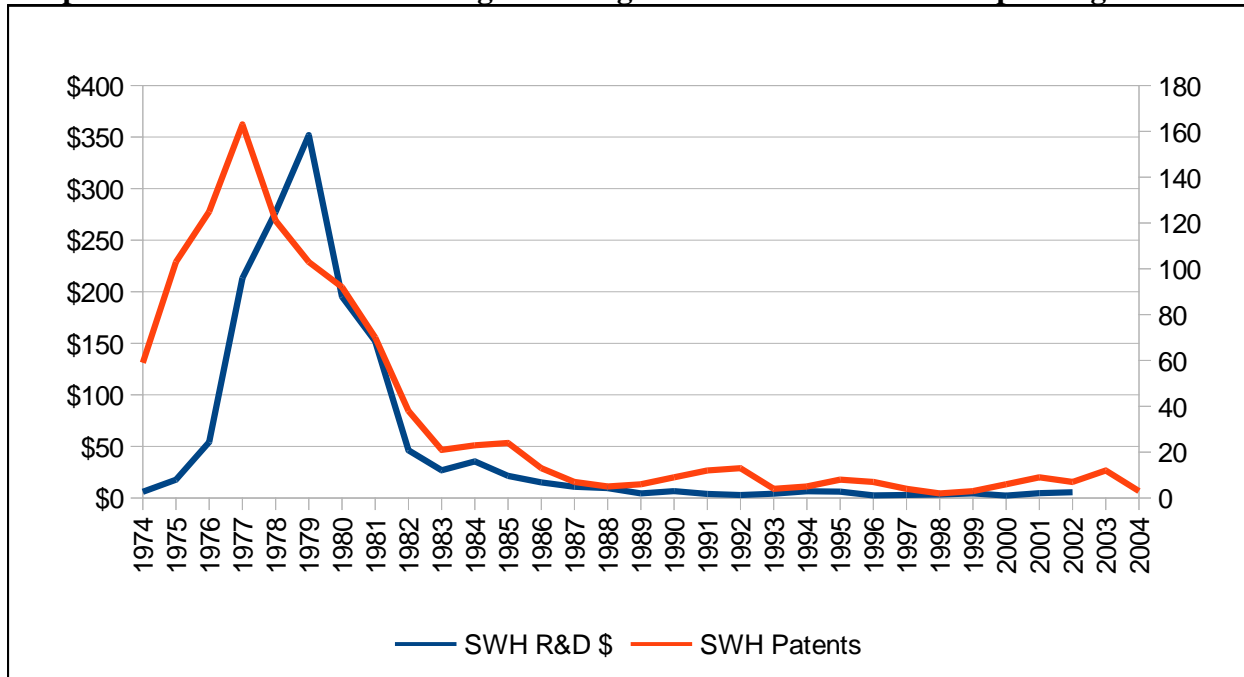
Unlike PV and STE, SWH did not experience a noticeable uptick in funding in the 1990s, after federal money fell off to near zero levels in the late 1980s. With its low cost profile and simple designs, there was simply not much innovation to fund.

Inventors have different reasons for filing (or not filing) patents, depending on their perception of the economic value of patents in their industry. In any technology-based industry targeted for patent analysis, it is important to try to understand this perception in order to place the results of analysis in context. In the SWH industry, the experts interviewed for this analysis had divergent opinions about whether patents covered the major innovations, with three saying they covered some of the major innovations, two saying they were unsure, and two saying that patents did not cover the major innovations. One response was that the "technology is not that sophisticated, so there is not much to patent" and another that there was lots of patenting in the 1970s because "in the '70s, the weirder it looked, the better it sold." One who answered that patents did cover the major innovations qualified it by saying "yes, but there was almost no innovation." (Taylor et al. 2007)

Neither funding nor patenting is likely to rise again for SWH: due to falling PV costs in recent years, it is now less costly to install a PV system and heat water with a heat pump (or even with resistance heat) than it is to install a SWH system (Holladay 2012, 2014). Moreover, with fewer moving parts and little to no freezing risk, PV will likely remain a more cost effective and less risky option for all but a few people in atypical circumstances.

4.2.3 Interviews

This section will detail four lightly structured interviews with experts in renewable-energy technology and emerging-technology financing. I addressed two primary questions to each of my interviewees, and then allowed a more free-form conversation to evolve from their initial answers. I also asked specifically about the role of government funds to the first two interviewees; however, this was not a relevant question to the latter two.

In this chapter's conclusion, I will use the qualitative information gleaned from these interviews to contextualize my analysis (in the manner of Borgonovi and O'hare 2004) of patenting within the four renewable-energy technologies addressed, and suggest why federal R&D money may have affected innovation in each technology differently.

Question 1: What does the R&D investment decision making process look like at your organization (or the organizations you work with)?

Question 2: What roll does invention, and patenting of inventions particular, play at your organization (or those you work with)? What about other forms of intellectual property protection?

Joseph Desmond – Senior VP of Marketing and Government Affairs, *BrightSource Energy*

Mr. Desmond has an extensive energy-related background, and I was particularly looking forward to the interview with him, since I imagined it would be highly informative and likely to

provide novel insights. He did not disappoint. Below is his biographical information from the BrightSource web site:

> *Joseph Desmond is Senior Vice President of Marketing and Government Affairs for BrightSource Energy. Desmond brings nearly three decades of private and public energy sector experience to his role at BrightSource Energy, where he oversees communications, marketing, and government and regulatory affairs. Prior to joining BrightSource Energy, Desmond served as Executive Vice President and Chief Marketing & Business Development Officer at Ice Energy, Inc. and Senior Vice President of External Affairs at NorthernStar Natural Gas. Desmond served in numerous executive roles under California Governor Arnold Schwarzenegger including Deputy Secretary of Energy for the State Resources Agency, Chairman of the California Energy Commission and Undersecretary for Energy Affairs. Prior to public service, Desmond spent four years as President and Chief Executive Officer of Infotility, Inc. Previously, he served as President and Chief Executive Officer of Electronic Lighting, Inc., and Vice President of Parke Industries. He serves on the Board of Directors for the American Council On Renewable Energy (ACORE). Desmond earned a B.S. in Marketing, Finance and Management from Northeastern University where he graduated magna cum laude.(BrightSource Energy 2017)*

Mr. Desmond stated that the vast majority of the innovation taking place at the utility level was process related, rather than more fundamental research. While many patents result from this type of research, most of them are on cost reduction techniques such as new installation methods, reducing number of components, improving assembly technique, use of alternative materials that decrease frequency of maintenance and replacement, and other process or design techniques that increase output. One example he gave was of the creation of a spacing techniques (using non-linear optimization algorithms) for the STE mirrors that increased energy output by between 7 and 12 percent. Another example was the application of wireless control technology to the mirrors to eliminate 85 percent of the cabling and related labor costs. Yet another was the development of a proprietary mirror coating that lasts four times as long as those of competitors.

He repeatedly emphasized that innovation at BrightSource and other firms like it was almost exclusively driven by cost considerations; however, decision about what R&D to invest in evolve as the company's business model evolves. He contrasted his business related experience with his time at the California Energy Commission: "I remember having to reorient myself to an R&D world where failure is an acceptable outcome. Failure is a research byproduct [in government] and that's okay." He indicated that many mid-sized firms don't have the option to invest in research that might fail, since it could easily bankrupt them. He also indicated that firms looked to industry consortia and government for leadership on the next generation of energy technologies, so as to predict where they would need to invest next, and to do the fundamental research that would support the industry that the individual firms could not afford to do.

Regarding intellectual property, he said, "Normally, when you have innovation, people want a way of validating it... [patenting] plays a role for convincing investors that are looking for protection and want to see that there is some sort of 'secret sauce' there... They do look at that to

see what is innovative, what is novel. Patents are one part of that. The other part is the industry knowledge, the experience learning, let's call them 'trade secrets' that you've learned by virtue of doing. They might not be patentable, but they are still highly valuable."

With regard to government funding, he said that while it did provide a resource, he had seen many companies engage in "hand wringing" about the prospect of needing to share revenue, pay royalties or fees, or share information with the government or public as a consequence of accepting public funds. In his experience, it was very rare for the funding agencies to actually exercise their right to payments, but it did provoke worry from businesses and served as a disincentive to accept government research money. Moreover, he said, "When you use public dollars, it's for the public good, and there is an expectation that you will disclose the results."

Daniel Purdy – Senior Technical Leader, *Electric Power Research Institute (EPRI)*

Mr. Purdy has been at EPRI for four years, and prior to that worked for GE, which provides him with varied perspectives on energy innovation. Our interview provided insights into the interface between the public and private sectors that EPRI creates, the role of EPRI in generating energy research in the public interest, and approaches to innovation unique to an organization like EPRI.

In discussing EPRI's role, he said, "We pool resources from the power generating industry and we work with [utilities] to solve communal problems. We prefer not to be a contracting engineering service to solve one particular plant's issue." About one third of their members are international and two thirds are domestic.

He indicated there was some relatively small participation from research labs and government as well. EPRI gets a small portion of their funding from the US federal government, but it receives most of its money from member utilities, so government funds have little influence on the direction of EPRI's research efforts. When dealing with government funds, "Like any company, we protect what we make, but there are clauses that make raw data available to the public." There are also typically medium-term exclusivity arrangements that protect intellectual property EPRI develops with government funds, but after approximately 5 years that IP is published by the government. He felt that patenting might not be a good metric of the kind of innovation that government funding promotes, given that much of it went to basic science.

To determine the focus of their research efforts, he said, "We examine the industry on the whole and ask a few questions: what's happening right now, what do we need to prepare for the future, and then what sort of grand trends – 20 year trends – can we help to predict?" To help gather this information, they hold meetings with members several times per year to understand what issues utilities are facing. Moreover, EPRI has a series of governing bodies, including the board of directors that is largely composed of CEOs of member utilities. "We get advice from a couple different levels about what our research should focus on, what our direction needs to be, what projects are important, down to what is important within a project and what our deliverables need to be."

EPRI has to do more fundamental research and think outside of the short-term time-frame that operators are forced into. "Ultimately we have the responsibility to look forward more than most of our colleagues running plants can afford to." About long-term or more basic research projects he said, "The plant's don't care. Most of those people aren't going to be in those jobs in

10 or 15 years, but the industry cares."

Mr. Purdy repeatedly emphasized that EPRI's work is in the public interest, and that that fact is central to how they work with organizations and choose what to research. He also indicated that a major role of EPRI was to help make connections between people in the electric power industry and facilitate knowledge transfer so as to improve the industry.

<u>Pieter Mul – Principal Consultant, Global Energy & Utilities, *PA Consulting Group*</u>

Mr. Mul analyzes energy markets, facilitates project finance and transactions for renewable installations (mainly wind) and advises clients on investments in energy firms. He noted an increased appetite for investment in renewables by his clients recently, and indicated that the growth of renewables into the mainstream was heavily influenced by both growing regulatory requirements (such as renewable portfolio standards) and the sharply falling costs of producing energy with the technology. Much of that he attributed to the ongoing investment tax credits. He sees the growth of distributed solar and utility scale solar as major trends in the near term electricity markets.

Much of the innovation funding that he sees in the market is coming from venture capital firms, and flowing to startups with an intriguing intellectual property. At the other end of the firm size scale, he noted R&D funding was also likely to come from large firms that can afford to invest in long-shot projects.

He also noted, "When most utilities build a power plant, it's in their interest for customers to use as much power as they can." Aligning incentives for utilities to invest in energy efficiency and clean energy is critical to fostering research in those areas.

<u>Meghan Burton – Corporate Lawyer, *Wilson Sonsini Goodrich & Rosati*</u>

While she had not been exposed to technology development and investment decision making processes at her clients' firms, Ms. Burton affirmed the central role that intellectual property plays for small firms seeking capital. She said, "Intellectual property is often the critical element in generating financing. The intellectual property is what is driving the value for the investors." She also affirmed that patenting was one of the principal ways that clients and their firms attempted to demonstrate to investors the value of their intellectual property – not only because of the indicia of originality and quality that patents convey, but also because a patent is an asset that can help mitigate risk for investors (i.e. if the company fails, the investors may be able to be paid back from the sale of the patent to a competitor or law firm). It also ensures the exclusivity of a firm's technology and assures investors that the firm will not be shut down by escalating licensing fees. These insights are confirmed empirically – firm value is positively associated with ownership of quality patents (Lanjouw and Schankerman 2004).

4.3   Principal Hypotheses

Since it is currently generally assumed that there is a causal relationship between public R&D spending and renewable energy patenting, similar to the relationships established for private firms by Hall, Griliches, and Hausman (1986), and Griliches (1990), my null hypothesis

will be that a significant relationship exists and that R&D will Granger-cause patents.

Alternatively, if the presence of increased patenting rates in a renewable energy technology serves to attract the attention of public entities making R&D investments, I may find that there exists a significant relationship between R&D and patents, with patents Granger-causing R&D spending. In this case, I might imagine that public R&D funding agencies end up lagging technology trends, only entering the field after the major burst of invention has already taken place. It is not hard to imagine why this might take place: instead of the fluid flow of information that takes place within a given firm, generating invention from research activity, government agencies may take some time to become aware of new technology trends and may also be reluctant to fund technology areas that have not shown some degree of prior inventive success.

Another possibility is that renewable energy patenting and public R&D are simply not significantly related. This could result from the type of research funded by public money being substantively different from inventive activity geared toward commercialization, where patenting is most relevant. Moreover, this could result from public R&D funding being indirectly causal of renewable energy patenting, such that necessary precursors to renewable energy inventions were generated by public R&D funding but in a remote enough way to obscure the modeled relationship.

4.4   Why This Matters / Is Interesting

In order to address issues of climate change, energy resource depletion, pollution, and other environmental and public health issues associated with the extraction and use of fossil fuels, commercially viable renewable energy technologies are necessary. In order to determine how to best support and expedite commercialization of renewable energy technologies, we need to know what policies have been successful at stimulating commercially directed invention in the past.

4.5   Data and Methods

To generate an empirical evaluation of the relationship between public R&D funding for renewable energy technologies and intensity of invention – as measured by patenting – I selected a statistical method that would allow for a variety of detectable relationships (vector auto-regression, or VAR) and would allow tests for Granger causality in order to determine the temporal direction of those relationships. Due to the importance of price signals in driving commercial invention, I also included a consumption weighted basket of fossil fuel prices as a control variable in one set of regressions. Further details about my methods are below.

4.5.1   Variable Construction

Patent Classes: Three of the four patent series were taken from data constructed by Taylor et al. (2007) for their report, including the solar PV, STE, and SWH patents. The solar PV patent series was constructed from a class-based search which was subsequently screened for irrelevant patents and coded. The STE and SWH patent series were constructed through abstract-based

keyword (Boolean) searches as described in Taylor et al. (2007) and subsequently screened and coded. The wind patent series was derived from a class-based search of patent class 290/55 (prime mover dynamo plants: wind), which was then screened and coded by myself and another research assistant (Victoria Fleming), yielding approximately 900 relevant wind-power related patents.

Public R&D expenditures: I used IEA national R&D expenditure data for the wind and three solar technologies, supplemented with material from Janet Sawin's dissertation for the 1992 and 1993 wind R&D data due to obvious errors in the IEA series. Note that although U.S. States are another source of public R&D support and aren't part of the national tally, the level of funds contributed by the States is small when compared to the federal levels. For example, of all 50 States, California is estimated to have provided by far the most significant level of public R&D money for solar technologies. California's records on these expenditures are spotty, but Taylor (2008a) pieces together credible R&D expenditures for three years in the relevant time period: Fiscal Year (FY)1977-78 (~3 million 2007$), FY 1981-82 (~1 million 2007$), and FY 1982-83 (~2 million 2007$). By contrast, federal solar R&D expenditures for similar years were much higher: 1977 (~435 million 2007$), 1981 (752 million 2007$), and 1982 (~343 million 2007$).

Fossil fuel prices: My fossil fuel price series is a consumption-weighted average of oil, coal, and natural gas. The price is expressed in year 2000 dollars per billion BTUs of energy, and the consumption weights are based on the annual percentage of each fuel's contribution to total fossil fuel fired electricity generation in the United States. Since the renewable energy technologies I consider generate electricity, I thought this would be the best weighting of fossil fuel prices to reflect incentives in the electricity generation market. Using total national fossil fuel consumption would overweight the contributions of oil and natural gas to prices.

4.5.2 Variable Testing

Given the propensity for time series data to have unit-roots, I first examined the data for non-stationarity, using Phillips-Perron tests with various lags (two through five, inclusive). I determined that the entire set of patent, R&D, and fossil fuel price series are non-stationary. Similar testing of the first differences of each series confirmed that each variable is I(1). Typically, cointegration and vector error correction models would be explored and developed when faced with a set of I(1) variables; however, for my purposes first differences are theoretically interesting. For example, modeling first differences of R&D and patents would explore the effect of changes in funding rates on changes in patenting rates, and/or vice versa.

4.5.3 Modeling Choice

In order to investigate these questions and test my hypotheses, I will employ various time-series methods including: tests of stationarity, selection order criteria tests, vector autoregression, various post-estimation tests (including tests of residual normality, complex eigenvector stability, and residual autocorrelation), and tests of Granger causality. To explore the potential for endogeneity in each of the patent and R&D variables, as well as the impacts of

multiple prior year changes on variables in the current year, I chose to use vector autoregression (VAR) modeling. VAR models treat each variable in turn as the dependent variable, regressing own-lags and other lagged variables in each model.

Due to the small number of years in my models, I use a small-sample size adjustment (resulting in reporting of small-sample t and F statistics for the models) as well as a small sample degrees-of-freedom adjustment (which includes a parameter adjustment when calculating the variance-covariance matrix of the errors). As shown in Table 3.3.4 below, I developed two sets of models. The first set uses first differences of patents and R&D for each of the technology classes (PV, STE, SWH, and Wind), as well as pooled patents and total federal renewable R&D funding. The second set also includes first differences of the consumption weighted basket of fossil fuel prices.

As shown in Tables 4.7.3.2.1 and 4.7.4.2.1, I developed two sets of models. The first set uses first differences of patents and R&D for each of the technology classes (PV, STE, SWH, and Wind), as well as pooled patents and total federal renewable R&D funding. The second set restricts the patents to those assigned to public entities or organizations working in collaboration with public entities.

4.5.4 Lag Structure Selection

Given the small sample size of the data and the degrees of freedom correction I make to the VAR models, the maximum possible lag structure eight years. When I tested the data using likelihood ratios, final prediction errors, and various information criteria (AIC, SBIC, and HQC) to determine the best lag structure, I found that the results were highly varying, indicating optimum maximum lags from zero to seven. I find the implication of zero to one lags implausible: in this case zero lags would imply no influence from prior years funding or patenting changes (as well as no ability to test for Granger causality), while only one lag would imply a very short-term impact from changes in patenting or R&D spending. These both seem highly unlikely since it often takes several years to generate research results (patents in this case) in any given technology, especially an emerging one.

Since I am not interested in conducting forecasting with the results of these models, but rather in exploring the historic relationship between R&D spending and patenting, I did not pre-select only one given lag structure. Instead, I examined the results of each model with maximum lags from two to seven in order to explore the characteristics of the models and consistency of their results. In addition to the significance of the models, I used tests of Granger causality and post-estimation testing (to detect assumption violations) in order to select models for presentation. I describe these tests and their results immediately below and in the results section.

4.5.5 Granger Causality

While the significance of the F-tests and individual coefficients in a VAR model reveals correlations between variables, these indicators fail to adequately answer the question of which variable could plausibly have cased the other. In order to explore possible causal relationships between the variables, I use the Granger causality test on each of the VAR models.

Granger causality does not necessarily imply true causality, but instead implies a

temporal relationship between correlated variables. This is accomplished with the use of the lagged terms in the VAR model. If the dependent variable is sufficiently correlated with lags of particular other variables, they are said to "Granger cause" the dependent variable. Granger causality may be detected due to true causality – e.g. variable X actually causes variable Y – or may be detected because of the relationship between modeled variables and an omitted variable (or several omitted variables) – e.g. omitted variable Z causes both variable X and Y, but changes in X occur before changes in Y; variable X causes variable Z, which in turn causes variable Y; or, variable Z causes variables X and W, and W then causes Y; etc. In addition, Granger causality may occur in both directions simultaneously – e.g. X Granger causes Y and Y Granger causes X. In this case, the strength of each relationship, as revealed by the degrees of statistical significance and the size of the standardized coefficients may be useful to compare.

While Granger causality is clearly limited in its ability to prove that one variable truly causes another, it is useful in ruling out possible causal relationships. While the results will show multiple cases of significant Granger causality, the lack of causality they also show is perhaps more interesting, as it cuts against one standard theoretical model of the relationship between R&D investment and patenting.

4.5.6   Post-Estimation Testing

I conducted multiple post-estimation tests on the VAR models in order to determine if any of the underlying assumptions were being violated, thereby impeaching the validity of the results. These tests included the Jarque-Bera test for normality of the error terms (which includes statistics for skewness, kurtosis, and the overall Jarque-Bera statistic), a Lagrange-multiplier test for autocorrelation in the residuals, and a test of the eigenvalue stability condition (testing that all of the eigenvalues of the VAR exist inside the complex unit-circle).

Of particular interest to me were the tests for autocorrelation and the stability conditions. Failure of either of these two tests caused me to seriously suspect the validity of a given model's results. While the Jarque-Bera test is also important, I am more liberal in consideration of models which have residuals that deviate slightly from normality (especially with regard to kurtosis). In several cases, mild but statistically significant kurtosis is present in a model's residuals. Upon inspection of histograms of the residuals, along with an overlaid kernel density graph and a normal distribution curve (see Appendix A), it is evident that the kurtosis represents only a very minor (although statistically significant) deviation from normality.

4.6   Model Specification

I chose a fairly simple VAR specification for my model for several reasons. First, hypothetical relationships between public R&D funding and patenting should be reasonably straightforward – either the funding is generating invention (i.e. patenting), rises in patenting rates are attracting the attention of public funding bodies and thereby generating R&D spending, or there is no relationship between public R&D and patenting.

Second, the lag structure of my models, coupled with the fairly small number of years in my data do not allow for the inclusion of many covariates. For example, with between three and six lags included, my usable data falls to between 22 and 28 years. This is at the boundary of

what will reliably generate significant results, even in the presence of a moderately strong relationship.

Therefore, I estimate a reduced form VAR equation (Equation 3, below), where y is the vector of endogenous variables at time t, and p is the maximum lag used in the model. Elements of y include public R&D spending on a particular technology type (or pooled spending) and patenting within a corresponding renewable energy technology (or pooled patenting for all technology types).

$$\Delta y_{m,t} = \alpha + \sum_{m=1}^{p} (\sum_{i=1}^{p} (\beta_{m,i} \Delta y_{m,t-i})) + \varepsilon_{m,t}$$

(3)

4.7   Results

4.7.1   Descriptive Statistics

Before presenting results from the VAR models, I will first explore some descriptive statistics regarding patents and public R&D for each technology. I will look at the distribution of government R&D funds, the distribution of patents with a government interest between technology classes, and the percentage of each set of technology class patents that has a government assignee or interest.

Comparing the relative levels of public R&D funding in each technology class and the relative distribution of patents with a government interest in each technology class reveals an interesting discrepancy: government R&D dollars are spread among technology classes much more evenly than patents with a government interest. See Table 4.7.1.1, below.

Table 4.7.1.1: Distribution of Government R&D and Patents with a Government Interest between Technology Classes

	Technology Classes			
	PV	STE	SWH	Wind
Percent of Gov. R&D in Each Tech. Class.	37.1%	26.6%	18.2%	18.2%
Percent of Pooled Gov. Int. Patents in Each Class	82.1%	4.1%	9.3%	4.5%

While total public funding for PV is roughly twice the level as for SWH or wind technologies, with STE funding falling in between, the number of patents in PV is 9 to 20 times the levels in STE, SWH, or wind. This suggests several possibilities: loose or idiosyncratic relationships between public funding levels and patenting, nonlinear relationships between public funding levels and patenting, or differing propensities to patent in different technologies,

among others. The percent of each technology class that is made up of patents with a government interest also varies considerably between technologies, with over 20 percent of all PV patents having at least some government interest, while between 5.2 and 6.8 percent of SWH, Wind, and STE patents have some government interest. This may indicate technology areas in which government R&D funding has played a larger role generating invention geared toward commercialization. See Table 4.7.1.2, below.

Table 4.7.1.2: Percent of Patents in Each Technology Class that have a Government Interest

	Technology Classes			
	PV	**STE**	**SWH**	**Wind**
Percent of Each Tech. Class that is Gov. Int.	20.1%	6.8%	5.2%	5.9%

4.7.2 VAR Model Lag Structures

The results of my post-estimation testing revealed violations of the stability conditions of the VAR estimates, autocorrelation of the residuals, and violations of the residual normality assumptions for some lag structures. I considered residual autocorrelation and stability violations to be more serious indicators of model misspecification, and discarded any models in which these violations occurred.

In the case of the normality assumptions, models for SWH and Wind yielded statistically significant kurtosis in the residuals; however, upon inspection of the errors in a histogram with an overlaid normal distribution curve and kernel density plot for the residuals, it is apparent that while there are statistically significant departures from normality, the substantive deviation is minimal and may be due entirely to my small sample size (see Appendix D, Graphs D1-D5). Therefore, I will present the results from these models. For a list of the full set of models presented in the following sections, see Table 4.7.2.1 below.

Table 4.7.2.1: VAR Models Selected with Lag Structure

	Technology Classes				
	Pooled Renewables	**PV**	**STE**	**SWH**	**Wind**
Dependent Variable(s)	R&D	R&D	Patents	Both Patents* and R&D*	Patents*
Lags Used	4	3	3	6	5

*Significant kurtosis was detected in the residuals of these models.

4.7.3 VAR Model Results for All Patent Series

4.7.3.1   Models with R&D Spending as the Dependent Variable

In addition to different lag structures, each model has a distinct pattern of statistically significant coefficients. Three of the models with R&D dollars as the dependent variable – pooled, PV, and SWH – exhibited statistically significant F tests and Granger-causality tests. The pooled model contained significant coefficients (at the less than one percent level) on the third and fourth lags on changes in patenting, and a significant (at the less than five percent level) own-lag on the third lag. The significant patenting lags had positive coefficients, indicating that increases in patenting three and four years prior, correlated with a positive change in pooled renewable R&D funding. The significant own-lag for R&D had a negative coefficient, indicating that increases in R&D funding three years prior were associated with decreases in current year funding, and vice versa. The R-squared value for the pooled R&D model was high, at 0.77, indicating a very strong fit, and the Granger-causality test was significant at the less than one percent level, indicating that patenting Granger-causes public R&D expenditures for the pooled model.

The PV model, with R&D as the dependent variable showed similar results to the pooled model. Significant coefficients were exhibited on two and three year lagged patents. The signs of the coefficients were positive, similar to the pooled model. The PV R&D model also had a highly significant F test (significant at the less than one percent level), a high R-squared value (0.68), and a highly significant Granger-causality test (significant at the less than one percent level), with patenting Granger-causing R&D funding. Due to the relatively large number of PV patents in the pooled model, these two models may both be reflecting the same underlying dynamic.

The SWH model, with R&D as the dependent variable differed from the other two, showing no significant patenting lags, while showing three significant own-lags, at two, four, and six years. Despite the lack of significant patenting variables, the F test and Granger-causality test were both highly significant (at the less than one percent level), indicating that an increase in SWH R&D Granger-causes an increase in funding two and six years later, and a decrease four years later. These results are highly suspicious and are further called into question when the R-squared value (0.99) is examined. Coupled with the kurtosis detected in the model, the extremely high R-squared value indicates overfitting of the model. This model should not be trusted and is likely due to omitted-variable bias and/or undetected autocorrelation.

4.7.3.2   Models with Patenting as the Dependent Variable

Two models with patenting of a renewable energy technology as the dependent variable (STE and SWH) exhibited significant Granger-causality tests, and two models (SWH and Wind) exhibited significant F tests.

The STE model, with patenting as the dependent variable, showed no significant own-lags and two significant R&D lags at two and three years, with one positive and one negative, respectively. The F test for this model was not significant, only reaching 0.16, however, the Granger-causality test was significant at the less than five percent level. The R-squared value

was 0.37, indicating a reasonable quality fit. Results for this model may be complicated by the small sample size or noise from non-government related patenting activity.

The SWH model, with patenting as the dependent variable, showed a significant own-lag at two years, and five significant lagged R&D variables, at one, two, three, five, and six years. The own-lag had a negative coefficient, indicating a decline in patenting levels two years after an increase. This may reflect cyclic invention patterns. The R&D lag coefficients were all positive, indicating positive impacts to patenting levels for many years after public investments were made. Also, the coefficients on the significant R&D lags decline monotonically, indicating a decreasing impact from more temporally distant R&D expenditure. Both F and Granger-causality tests were significant for this model; however, the R-squared value was quite high (0.95), which again indicates the possibility of overfitting and/or effects of a small sample size. More targeted research on the history of solar water heating invention is needed before definitive judgments can be made.

The wind model, with patenting as the dependent variable, showed significant own lags at one and two years, with negative and positive coefficients, respectively. Again, these may be indicative of some cyclic pattern in wind turbine invention. No significant lags of public R&D spending existed. The F test was highly significant, and the R-squared value (0.75) indicated a good model fit. Kurtosis was again noted with this model. The Granger-causality test, with R&D Granger-causing patents, was not significant at the five percent level, although it very nearly was.

See Table 4.7.3.2.1 for these VAR model results, below.

Table 4.7.3.2.1: VAR Models – Public R&D Expenditures and Renewable Energy Technology Patents

Dep. Var.=	Pooled Ren. R&D	(SE)	PV R&D	(SE)	SWH R&D	(SE)	STE Patents	(SE)	SWH Patents	(SE)	Wind Patents	(SE)
Regression Covariates	n=26		n=25		n=22		n=25		n=22		n=25	
Patents												
Lag = 1	-0.37	0.88	-0.26	0.40	-0.06	0.40	0.12	0.20	-0.19	0.29	-0.64	0.26**
Lag = 2	0.24	0.89	1.09	0.46**	-0.39	0.39	-0.03	0.22	-0.76	0.28**	0.68	0.27**
Lag = 3	4.18	0.81***	1.82	0.47***	0.64	0.37	-0.07	0.16	-0.27	0.27	0.60	0.31*
Lag = 4	3.44	1.03***	--	--	0.01	0.37	--	--	-0.42	0.27	0.08	0.29
Lag = 5	--	--	--	--	-0.33	0.27	--	--	-0.40	0.20*	-0.06	0.18
Lag = 6	--	--	--	--	-0.26	0.22	--	--	-0.16	0.16	--	--
R&D												
Lag = 1	-0.15	0.19	0.18	0.17	0.12	0.21	-0.04	0.02	0.47	0.15***	0.05	0.04
Lag = 2	-0.12	0.15	-0.31	0.17*	0.55	0.21**	0.08	0.03**	0.44	0.15**	0.09	0.04*
Lag = 3	-0.40	0.16**	-0.13	0.16	0.03	0.10	-0.06	0.02**	0.16	0.07**	-0.04	0.05
Lag = 4	-0.22	0.16	--	--	-0.22	0.06***	--	--	0.04	0.05	-0.10	0.05*
Lag = 5	--	--	--	--	0.10	0.08	--	--	0.14	0.06**	-0.05	0.06
Lag = 6	--	--	--	--	0.16	0.06**	--	--	0.13	0.04**	--	--
Constant	-2.20	22.18	1.17	6.70	0.33	1.39	-0.28	0.98	-0.50	1.01	1.00	0.92
P>F	0.00***		0.00***		0.00***		0.16		0.00***		0.01***	
R-Squared	0.77		0.68		0.99		0.37		0.95		0.75	
Granger-Causality	0.00***		0.00***		0.00***		0.04**		0.02**		0.05*	

*significant at the <10% level **sig. at the <5% level ***sig. at the <1% level

4.7.4 VAR Model Results for Patents with a Government Interest

The relationship between public R&D and overall patenting may be complicated by the high prevalence of fully private patents in the renewable-energy-technology patent counts. One might offer an explanation of the idiosyncratic relationship between patenting and public R&D by hypothesizing that while private patents sometimes precede public funding, public funding will still generate public patenting; public R&D should Granger-cause government interest or assigned patents, and the reason I have not detected this is due to the "noise" from private patenting, and/or due to private patenting stimulating government R&D funding, which in turn stimulates government interest patenting.

Consequently, I have also run the same set of models using annual counts for government interest or assigned patents in each technology class. The selection order criteria indicated different optimal lag structure from the previous set of models, shown in Table 4.7.4.1, below.

Table 4.7.4.1: VAR Models Selected with Lag Structure (Gov. Int. Patents)

	Technology Classes				
	Pooled Renewables	**PV**	**STE**	**SWH**	**Wind**
Dependent Variable(s)	R&D	R&D*	Both Patents** and R&D	Both Patents and R&D*	Patents
Lags Used	2	5	2	4	2

*Significant kurtosis was detected in the residuals of these models.
**Significant skewness was detected in the residuals of this model.

4.7.4.1   Models with R&D Spending as the Dependent Variable

In addition to different lag structures, each model has a distinct pattern of statistically significant coefficients. Four of the models with R&D dollars as the dependent variable – pooled, PV, STE, and SWH – exhibited statistically significant F tests, while three – PV, STE, and SWH – had significant Granger-causality tests. The pooled model contained one significant coefficients (at the less than one percent level) on the first own lag. The significant R&D first lag had positive coefficient, indicating that a change in prior year pooled funding correlated with a change of similar direction in pooled current year funding. The R-squared value for the pooled R&D model was reasonably high, at 0.38, and the F test was significant at the less than 5 percent level; however, the Granger-causality test was not significant, reflecting the general weakness in the model.

The PV model, with R&D as the dependent variable showed highly significant patent lags and own lags. Significant coefficients were exhibited on three and five year lagged patents, both with large coefficients, indicating same direction changes in funding at those two intervals after changes in patenting rates. The own lags were significant at two and five years, with both coefficients being negative, indicating opposite direction changes at those intervals. The PV

R&D model also had a significant F test (significant at the less than five percent level), a high R-squared value (0.76), and a highly significant Granger-causality test (significant at the less than one percent level).

The STE model, with R&D as the dependent variable showed a significant patenting first lag and a significant first own lag. Both were highly significant (at the less than one percent level) with positive coefficients. The F test was also highly significant, while the Granger-causality test was significant at the less than 5 percent level. The R-squared value was 0.47, indicating a good model fit.

The SWH model, with R&D as the dependent variable showed significant patenting lags at two and four years (both with large and highly significant coefficients), and three significant own lags, at two, three, and four years. The two-year own lag had a positive coefficient, while the three and four-year lags had negative ones, all of which were highly significant. The F test and Granger-causality test were both highly significant (at the less than one percent level), indicating that both patenting and R&D funding changes Granger-cause SWH R&D. The R-squared value was 0.84, indicating an excellent model fit; however, the presence of significant kurtosis in the model indicates possible over-fitting, which is consistent with the findings for the all SWH patents model presented in the prior set of regressions. It seems as though eliminating the purely private-interest patents from the SWH series reduced the possible problems with modeling SWH patents, but may not have fully eliminated them.

4.7.4.2   Models with Patenting as the Dependent Variable

Two models with patenting of a renewable energy technology as the dependent variable (STE and Wind) exhibited significant Granger-causality tests, and three models (STE, SWH, and Wind) exhibited significant F tests.

The STE model with government-interest patenting as the dependent variable, showed a highly significant first own lag, with a negative coefficient, and no significant lags on R&D funding. The F test for this model was highly significant, and the R-squared value was 0.55; however, the Granger-causality test was not even close to significant. These results are almost exactly opposite those of the all-patents STE model presented above. That, coupled with the skewness detected in the model indicates its weakness.

The SWH model with government-interest patenting as the dependent variable, showed a significant first own lag, and two significant lagged R&D variables, at one and two years. The own-lag had a negative coefficient, indicating a decline in patenting levels two years after an increase. This may reflect cyclic invention patterns. The R&D lag coefficients were both positive, indicating positive impacts to patenting levels for several years after public investments were made. Also, the coefficients on the significant R&D lags decline monotonically, indicating a decreasing impact from more temporally distant R&D expenditure. Both F and Granger-causality tests were highly significant for this model; and, the R-squared value was quite high (0.82). The high R-squared value again indicates the possibility of over-fitting and/or effects of a small sample size, however it's not as egregious as in the all-patents SWH model presented above, indicating that eliminating the private-interest patents may have improved the model quality. No skewness or kurtosis was present in the residuals for this model, which is promising.

The wind model with government-interest patenting as the dependent variable, showed a

51

significant first own lag with a negative coefficient, and a significant first lag on R&D funding with a positive coefficient (the two year lag was nearly significant as well, and had a smaller positive coefficient). This is only the third model to what seems to be an impact from R&D funding changes to patenting, with that impact tapering off over time. The F test was highly significant, and the R-squared value (0.72) indicated a good model fit (and was similar to the all-patents wind model presented above). Unlike the all-patents wind model, kurtosis was not present in model residuals, indicating that elimination of private-interest patents may have solved some problems. The Granger-causality test was highly significant.

See Table 4.7.4.2.1 for these VAR model results, below.

Table 4.7.4.2.1: VAR Models – Public R&D Expenditures and Government Interest or Assigned Renewable Energy Technology Patents

Dep. Var.=	Pooled Ren. R&D	(SE)	PV R&D	(SE)	STE R&D	(SE)	SWH R&D	(SE)	STE Patents	(SE)	SWH Patents	(SE)	Wind Patents	(SE)
Regression Covariates	n=28		n=22		n=26		n=24		n=26		n=24		n=28	
Patents														
Lag = 1	6.50	4.62	3.45	1.93	14.43	4.98***	-3.38	3.27	-0.93	0.20***	-0.38	0.16**	-1.15	0.24***
Lag = 2	-1.17	4.77	0.43	2.24	-2.43	8.379	14.72	3.66***	-0.20	0.33	-0.29	0.18	-0.33	0.20
Lag = 3	--	--	9.00	2.05***	--	--	5.56	3.63	--	--	-0.01	0.18	--	--
Lag = 4	--	--	0.25	2.18	--	--	9.12	2.98***	--	--	0.09	0.15	--	--
Lag = 5	--	--	6.79	1.74**	--	--	--	--	--	--	--	--	--	--
R&D														
Lag = 1	0.60	0.21***	0.16	0.21	0.71	0.24***	0.04	0.18	-0.01	0.01	0.03	0.01***	0.04	0.01***
Lag = 2	-0.17	0.20	-0.70	0.22***	-0.26	0.23	0.40	0.11***	0.01	0.01	0.01	0.005**	0.03	0.01*
Lag = 3	--	--	-0.37	0.23	--	--	-0.55	0.11***	--	--	0.00	0.01	--	--
Lag = 4	--	--	-0.15	0.22	--	--	-0.45	0.15***	--	--	0.00	0.01	--	--
Lag = 5	--	--	-0.57	0.19***	--	--	--	--	--	--	--	--	--	--
Constant	2.67	32.13	-2.51	6.99	-3.50	7.15	-5.07	4.81	0.03	0.28	0.01	0.24	0.15	0.23
P>F	0.02**		0.02**		0.01***		0.00***		0.00***		0.00***		0.00***	
R-Squared	0.38		0.76		0.47		0.84		0.55		0.82		0.72	
Granger-Causality	0.36		0.01***		0.01**		0.00***		0.49		0.00***		0.00***	

*significant at the <10% level **sig. at the <5% level ***sig. at the <1% level

4.7.5 Fossil Fuel Price Models

I also used a consumption weighted basket of fossil fuel prices as an exogenous covariate in a set of VAR regressions, with contemporaneous prices as well as various lags included. Fossil fuels were not included as an endogenous variable as I deemed it unreasonable to expect renewable energy patenting or research funding to influence their prices in a meaningful way, especially given the small size of the renewable energy market.

The inclusion of lagged fossil fuel prices resulted in occasional significant coefficients; however, these results did not alter the relationships between R&D and patents, they did not alter the Granger-causality conclusions of my presented models, and they offered little in the way of meaningful insight (e.g. a small but significant negative coefficient on a three-year lag of fossil fuel prices, indicating that patenting or R&D would slightly decline three years after a rise in fossil fuel prices) while decreasing the power of my regressions. Therefore, I have not included these results here.

4.8 Conclusions

This paper examines the relationship between renewable energy invention, as indicated by patenting rates, and publicly funded R&D spending from 1974 to 2008 in the United States. It is an attempt to apply some methodological rigor to casual and frequent claims (both popular and scholarly) that there is a strong relationship between public research funding and innovation in renewable energy technologies. Complex policy details and nuances have typically been overlooked when discussing public research funding in broad terms, and my analysis here demonstrates the importance of attending to such details, since the relationships of patenting rates and public R&D funding differ widely depending on the technology and funding program.

The twenty models analyzed above (ten for all patents and ten for government interest patents only) describe a variety of relationships between R&D funding and patenting, relationships that may be indicative of fine grained policy detail differences within each funding program, and different incentive structures within each industry. To be clear, I do not believe that these models should be used to predict effects (or lack of effects) from public funding of various renewable energy R&D technologies. The kurtosis present in several of the models' residuals alone would indicate caution in that regard. Instead, I think they are illustrative of the idiosyncratic historical relationship between public R&D expenditures and invention (as measured by patenting) in these different technologies, and indicative of the vital role for careful policy design when attempting to motivate invention with an eye toward commercialization.

In fact, my diverse modeling outcomes – from R&D spending being Granger-caused by patenting, to patents being Granger-caused by R&D spending, to patents and R&D spending Granger-causing each other – indicates that policy design details are vitally important in the case of government funding for renewable energy R&D. This is also suggested by the changes in modeling results between the all-patent models and the government-interest models – the isolation of government-interest patents reveals a more pervasive situation of patenting Granger-causing R&D funding than was seen in the all-patents models. Incentive structures, institutional relationships, geography, and methods of knowledge sharing within a given funding program may determine if public R&D funding is successful in actually promoting targeted invention, or

merely acts as a response to newly popular technology areas. This result may be heartening to policy analysts, and implementation specialists.

Regardless of the causal directions, my results do not mean that we should end public R&D expenditures; it is quite possible that the non-patented R&D activities of public entities are extremely valuable in other ways. However, it means that we need to develop a more sophisticated understanding of the direct and indirect impacts of R&D funding. The monopoly rights of patents are granted to reward invention, but patents also exist to grow publicly available knowledge on new technical advances as quickly as possible. Public R&D expenditures might be better thought of as contributing to that knowledge stock in the long term, or contributing in other ways than through an influence on patents, ways that are less directly measurable.

Interviews with energy experts seem to confirm that public R&D expenditures are more valuable in generating fundamental science, exploring speculative technologies, and/or helping to predict and promote the long-term evolution of the energy industry. The types of cost-reducing innovations that are the focus of R&D by small and mid-sized firms are typically proprietary, and sometimes patented mainly to signal value to investors. Moreover, some firms actively engaged in commercializing renewable-energy are reluctant to accept public R&D funding for fear of having to disclose results to the public.

Therefore, given my results, using public R&D dollars to motivate increased commercialization of existing technologies may be misguided. Instead, given the role of public R&D in the creation of basic research, foundational technologies, and serendipitous discoveries (Popp 2006, Joint Economic Committee 2010), such funds might better be spent on long-shot technologies that are not currently commercially viable. Future research on the success of public R&D seeding such technologies would be valuable.

For the purposes of motivating near-term affordability of established technologies – ones that already have some demonstrated functionality, but are not yet cost competitive with conventional energy sources – different approaches are likely needed. Renewable portfolio standards, feed-in tariffs, and pollution fees are other policy options that have demonstrated success (J. Desmond, phone interview, April 25, 2017; Mitchell, Bauknecht, and Connor 2006; Butler and Neuhoff 2008; Johnstone, Hascic, and Popp 2010), are easily linked to growth in renewable technology implementation, and may thereby generate learning-effect benefits.

Alternatively, funds could be spent directly on the technologies themselves (e.g. purchase and installation of solar PV systems for government buildings) or on support for private entities purchasing the technologies (e.g. tax credits for installing renewable systems). Public funding for commercialization-oriented renewable-energy invention seems to be a less certain method.

5 Overall Conclusions and Implications

Theory is cheap.
― Unknown origin

5.1 Introduction

This chapter will review the major findings from the research done in the prior two chapters, discuss them in the context of the broader innovation literature from chapter 2, and then draw out implications for future policy making and policy research (specifically in the area of technology policy).

The previous two chapters share a unifying theme: together, they attempt to generate and then apply high-quality methods to questions of public funding of R&D activities in young technology sectors. As mentioned in the overview section, this theme has policy relevance not merely because it addresses programs receiving public funds, but also because of the potential for those technologies to generate public benefits and positive externalities.

This research contributes to the policy literature on innovation, specifically to the research on patents as a metric of invention, by challenging conventional and potentially spurious causal narratives and the methods used to affirm them. Moreover, it brings to bear improved methods on all aspects of patent data analysis, from the selection and coding of the data, to the statistical techniques used to draw inference.

5.2 Overall Findings

This dissertation has two major findings, which can be summarized as: 1) several of the statistical methods regularly used to analyze and draw inference from patent data are prone to extremely high rates of spurious correlations when used on non-stationary time-series, and since patent data appear to be difference-stationary series, these methods should be avoided; and, 2) when using differenced patent-data to analyze the (Granger-causal) relationships between federal R&D funding and renewable-energy patenting, a more complex set of interactions becomes evident than has been typically found using the aforementioned error-prone methods; these interactions imply that policy details matter when trying to incentivize innovation. We cannot accept the standard narrative that increased federal R&D spending will generate needed energy-technology patenting (a proxy for invention) without better evidence.

Moreover, it's far from clear that federal R&D funding *ought* to be generating increased renewable-energy patenting. Energy experts I interviewed consistently emphasized the use of patents by private firms to signal value to investors and the market. Conversely, federal R&D money is intended to be used for advancing the public interest, often focuses on fundamental research rather than commercialization, and some firms are reluctant to accept federal money for fear of having to share their discoveries. All other things being equal, increased research funding should generate *some* amount of increased patenting, but it might be a weak enough signal so as to be lost in the noise.

Scholarly articles noted in chapter 4, and countless popular articles about renewable energy, overlook complex policy details and nuances when discussing public research funding,

and my analysis demonstrates the importance of attending to such details, since the relationships of patenting rates and public R&D funding differ widely depending on the technology and funding program.

For the purposes of motivating near-term affordability of functioning technologies – ones that already have some demonstrated capabilities, but are not yet cost competitive with conventional energy sources – different approaches are needed. Public funding for commercialization oriented renewable-energy invention seems to be a less certain method, at least with the types of funding programs that have existed so far.

As Janet Sawin noted in her exhaustive study of wind-power programs in the United States and Europe:

> *The future of wind energy and other alternative technologies will rely primarily on government policy. Above all, the success of these technologies – the rate and level of their development and diffusion – is a matter of policy choice. In order to be adopted on a meaningful scale, they will require effective, appropriate and consistent policies that are flexible, forward looking, and legislated with a long-term view toward advancing a technology and an industry. Simply throwing money at technologies – for research and development, for example – is not the answer.* (Sawin 2001)

Finally, the capabilities of many of the policy instruments detailed in chapter 2 provide viable alternatives for supporting the development of renewable-energy. A variety of government actions can stimulate technological innovation, including the provision of positive inducements, such as tax breaks, contracts, and prizes; the facilitation of knowledge sharing; and regulation or outright prohibition of certain methods.

With regard to regulation, a study of multiple factors influencing environmental innovation found that "government regulation appears to be a greater stimulus to inventive activity than government-sponsored research support alone, and that the anticipation of regulation also spurs inventive activity." Firms, being strategic, will innovate in anticipation of new regulations in an attempt to not only be undamaged by the regulation, but also to gain market advantage over their competitors when the regulations are put in place, and thereby increase their profits. In addition, the strictness of the regulation induces concentration along particular technology paths, thereby not only influencing quantity but also type of innovation. Finally, significant impacts occur from government sponsored knowledge sharing activities, such as technical conferences (Taylor et al. 2003).

Regulation, especially in the environmental industry, acts to secure demand for innovations along particular technology paths, and thereby assure firms of market opportunities, which is a strong incentive to innovate (Mowery and Rosenberg 1982). In this way, regulation can be seen to play a somewhat similar role to positive government inducements to innovate, as it indirectly shifts firms' profit probability distributions; however, wasteful effort is less likely than in the case of subsidies and tax incentives since unsuccessful innovations are not rewarded.

5.3   Implications for Policy Research

Taken together, my findings demonstrate the interdependent relationship between appropriate analytic techniques and accurate analysis when examining the sometimes subtle effects of complex policies. If social science research is to have any real value, if it is to do more than merely make us feel secure in our own knowledge, if it is to provide real guidance to policy making, it must generate predictions with a greater probability of being accurate than random guesses. Theory may be cheap (or more accurately, cheap theory is always available), but so too may be empirical analysis. The enormous problem of unreproducible results and questionable methods is slowly being confronted by the research community, and much of the patent-data research needs to be reexamined in light of the findings above.

Interviews

Burton, Meghan (2017, April 27). Wilson Sonsini Goodrich & Rosati. Phone interview.

Desmond, Joseph (2017, April 25). BrightSource Energy, Inc.. Phone interview.

Mul, Pieter (2017, April 27). PA Consulting Group. Phone interview.

Purdy, Daniel (2017, April 26). Electric Power Research Institute. Phone interview.

Bibliography

Abernathy, W. J. (1978). *The Productivity Dilemma.* Baltimore, MD; Johns Hopkins University Press.

Abernathy, W. J., & Clark, K. B. (1985). Innovation: mapping the winds of creative destruction. *Research Policy, 14 (1),* 3-22.

Abernathy, W. J., & Utterback, J. M. (1978). Patterns of innovation in technology. *Technology review*, *80*(7), 40-47.

Adams, J., & Griliches, Z. (1996). Measuring science: an exploration. *Proceedings of the National Academy of Sciences, 93,* 12664-12670.

Anderson, P. & Tushman, M. L. (1990). Technological Discontinuities and Dominant Designs: A Cyclical Model of Technological Change. *Administrative Science Quarterly,* 35(4): 604-633.

Archibugi, D. (1988). In search of a useful measure of technological innovation (to make economists happy without discontenting technologists). *Technological Forecasting and Social Change, 34 (3)*, 253-77.

Archibugi, D., & Pianta, M. (1996). Measuring technological change through patents and innovation surveys. *Technovation, 16 (9),* 451-68.

Arrow, K. (1962). The Economic Implications of Learning by Doing. *Review of Economic Studies, 29*, 155-173.

Arundel, A., & Kabla, I. (1998). What percentage of innovations are patented? Empirical estimates for European firms. *Research Policy, 27,* 127-141.

Arundel, A., & Kemp, R. (2009). Measuring eco-innovation. *UNU-M ERIT Working Paper Series, 2009-017.*

Ashford, N.A., Ayers, C., & Stone, R.F. (1985). Using regulation to change the market for innovation. *Harvard Environmental Law Review, 9*, 419-66.

Audretsch, D. B., & Feldman, M. P. (1996). R&D Spillovers and the geography of innovation and production. *The American Economic Review, 86 (3),* 630-640.

Baptista, R., & Swann, P. (1998). Do firms in clusters innovate more?. *Research policy*, *27*(5), 525-540.

Barras, R. (1984). Towards a theory of innovation in services. *Research Policy, 15*, 161–73.

Basberg, B.L. (1987). Patents and the measurement of technological change: A survey of the literature. *Research Policy, 16 (2-4),* 131-41.

Beck, N., & Katz, J. N. (2011). Modeling dynamics in time-series–cross-section political economy data. *Annual Review of Political Science*, *14*, 331-352.

Becker, H. S. (2014). *What about Mozart? What about murder?: Reasoning from cases*. University of Chicago Press.

Becker, W. E., & Kennedy, P. E. (2005). Does teaching enhance research in economics?. *The American economic review*, *95*(2), 172-176.

Benner, M. J., & Tushman, M. (2002). Process management and technological innovation: A longitudinal study of the photography and paint industries. *Administrative science quarterly*, *47*(4), 676-707.

Block, D., & Harrison, J. (1997). *Solar water heating: A question and answer primer*. Florida Solar Energy Center.

Blundell, R., Griffith, R., & Van Reenen, J. (1995). Dynamic count data models of technological innovation. *The Economic Journal*, 333-344.

Borgonovi, F., & O'hare, M. (2004). The impact of the national endowment for the arts in the United States: Institutional and sectoral effects on private funding. *Journal of cultural Economics*, *28*(1), 21-36.

Brandt, P. T., Williams, J. T., & Fordham, B. (1998, July). Modeling time series count data: A state-space approach to event counts. In *prepared for the 1998 Society for Political Methodology Meeting, University of California, San Diego*.

Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'Engines of growth'?. *Journal of econometrics*, *65*(1), 83-108.

Brunnermeier, S. B., & Cohen, M. A. (2003). Determinants of environmental innovation in US manufacturing industries. *Journal of environmental economics and management*, *45*(2), 278-293.

Butler, L., & Neuhoff, K. (2008) Comparison of feed-in tariff, quota and auction mechanisms to support wind power development. *Renewable Energy*, 33 (8), 1854-1867

California Energy Commission (CEC). (1981). *California Wind Directory* (7), Sacramento CA.

Cameron, A. C., & Trivedi, P. K. (1990). Regression-based tests for overdispersion in the Poisson model. *Journal of Econometrics*, *46*(3), 347-364.

Carayol, N., & Matt, M. (2004). Does research organization influence academic production? Laboratory level evidence from a large European university. *Research Policy, 33,* 1081-1102.

Carayol, N., & Matt, M. (2006). Individual and collective determinants of academic scientists' productivity. *Information Economics and Policy, 18,* 55-72.

Carlaw, K. I., & Lipsey, R. G. (2002). Externalities, technological complementarities and sustained economic growth. *Research Policy*, *31*(8), 1305-1315.

Carpenter, M.B., Narin, F., & Woolf, P. (1981). Citation rates to technologically important patents. *World Patent Information,* 160-3.

Chen, C. J., & Huang, J. W. (2009). Strategic human resource practices and innovation performance—The mediating role of knowledge management capacity. *Journal of business research*, *62*(1), 104-114.

Chen, H., Roco, M., Li, X., & Lin, Y. (2008). Trends in Nanotechnology Patents. *Nature Nanotechnology.* 3:123-5

Clarke, N., & Riba, M. (1998). *Patent information for technology foresight*. Vienna, Austria: European Patent Office.

Coe, D. T., & Helpman, E. (1995). International R&D spillovers. *European Economic Review, 39,* 859-887.

Cohen, W., & Levin, R. (1989). Empirical studies of innovation and market structure. In R. Schmalensee, & R.D. Willig (Eds.), *Handbook of industrial organization* (Vol. 2, pp. 1059-1107). Amsterdam: Elsevier.

Cohen, W., & Levinthal, D. (1989). Innovation and learning: the two faces of R&D. *The Economic Journal, 99 (397),* 569-596.

Cohen, W., & Levinthal, D. (1990). Absorptive capacity: a new perspective on learning and innovation. *Administrative Science Quarterly, 35 (1),* 128-152.

Cohen, W., Nelson, R., & Walsh, J. (2001). Protecting their intellectual assets: appropriability conditions and why firms patent and why they do not in the American manufacturing sector. *NBER Working Paper 7552*.

Cohen, W., Nelson, R., & Walsh, J. (2002). Links and impacts: the influence of public research on industrial R&D. *Management Science, 48,* 1-23.

Damanpour, F. (1991). Organizational innovation: a meta-analysis of the effects of determinants and moderators. *The Academy of Management Journal, 34 (3),* 555-590.

Dang, Y., Zhang, Y., Hu, P., Brown, S., & Chen, H. (2010). Knowledge mapping for rapidly evolving domains: A design science approach. *Decision Support Systems*, 50: 415-27

Devinney, T. (1987). Entry and Learning. *Management Science, 33 (6),* 706-724

Dewar, R., & Dutton, J. (1986). The adoption of radical and incremental innovations: an empirical analysis. *Management Science, 32 (11),* 1422-1433.

Dickey, D. A., Bell, W. R., & Miller, R. B. (1986). Unit roots in time series models: Tests and implications. *The American Statistician*, *40*(1), 12-26.

Dooley, J. J. (2008). US federal investments in energy R&D: 1961-2008. *Pacific Northwest National Laboratory. www. pnl. gov/main/publications/external/technical_reports/PNNL-17952. pdf*.

Dougherty, D. (1992). Interpretive barriers to successful product innovation in large firms. *Organizational Science, 3 (2),* 179-202.

Dupuy, D. (1997). Technological change and environmental policy – The diffusion of environmental technology. *Growth and Change, 28,* 49-66.

Energy Reorganization Act of 1974. (ERA). (1974). Public Law 438, 93rd Congress, second session.

Erkal, N., & Scotchmer, S. (2009). Scarcity of ideas and R&D options: use it, lose it or bank it. *NBER Working Paper 14940.*

Fagerberg, J., & Verspagen, B. (2002). Technology-gaps, innovation-diffusion and transformation: an evolutionary interpretation. *Research policy*, *31*(8), 1291-1304.

Farrell, J., & Saloner, G. (1985). Standardization, compatibility, and innovation. *The RAND Journal of Economics*, 70-83.

Finardi, U. (2011). Time relations between scientific production and patenting of knowledge: the case of nanotechnologies. *Scientometrics,* 89: 37-50

Giersch, H. (Ed.). (1982). *Emerging technologies: consequences for economic growth, structural change, and employment: Symposium 1981*. Mohr.

Goolsbee, A. (1998). Does government R&D policy mainly benefit scientists and engineers? *American Economic Review, 88 (2),* 298–302.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424-438.

Granger, C. W., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of econometrics*, *2*(2), 111-120.

Greiner, A., & Semmler, W. (2000). Endogenous growth, government debt and budgetary regimes. *Journal of Macroeconomics*, *22*(3), 363-384.

Griliches, Z. (1979). Issues in assessing the contribution of research and development to productivity growth. *The Bell Journal of Economics, 10 (1),* 92-116.

Griliches, Z. (1990). Patent statistics as economic indicators: a survey. *The Journal of Economic Literature, 28 (4),* 1661-1707.

Griliches, Z. (1994). Productivity, R&D, and the data constraint. *The American Economic Review, 84 (1)*, 1-23.

Guo, G. (1996). Negative Multinomial Regression Model for Clustered Event Counts. *Sociological Methodology,* 26: 113-132.

Hahn, F., & Solow, R. M. (1997). *A critical essay on modern macroeconomic theory*. MIT press.

Hall, B. H. (1996). The private and social returns to research and development. *Technology, R&D, and the Economy*, *140*, 162.

Hall, B. (2004). Exploring the patent explosion. *NBER Working Paper 10605*.

Hall, B., Griliches, Z., & Hausman, J.A. (1986). Patents and R&D: Is there a lag? *International Economic Review, 27 (2),* 265-83.

Hall, B. H., Jaffe, A. B. & Trajtenberg, M. (2000). *Market Value and Patent Citations: A First Look*. Cambridge, MA, NBER Working Paper W7741.

Hansen, J.A. (1992). *New indicators of industrial innovation in six countries: A comparative analysis.* Washington, D.C.: National Science Foundation.

Heaton, G. R. (1990). Regulation and technological change: charting a new emphasis. In *World Resources Institute symposium" Toward* 2000: Environment, Technology, and the New Century (pp. 87-89).

Heller, M. A., Eisenberg, R. S. (1998). Can patents deter innovation? The anticommons in biomedical research. *Science, 280,* 698-701.

Holladay, M. (2012, March 12). Solar Thermal Is Dead [Web log post]. Retrieved April 13, 2017 from http://www.greenbuildingadvisor.com/blogs/dept/musings/solar-thermal-dead

Holladay, M. (2014, December 26). Solar Thermal Is Really, Really Dead [Web log post]. Retrieved April 13, 2017 from http://www.greenbuildingadvisor.com/blogs/dept/musings/solar-thermal-really-really-dead

Huang, Z., Chen, H., Yan, L., & Roco, M. (2005). Longitudinal nanotechnology development (1991–2002): National Science Foundation funding and its impact on patents, *Journal of Nanoparticle Research.* 7: 343-76

Jaffe, A. B. (1996). Trends and patterns in research and development expenditures in the United States. *Proceedings of the National Academy of Sciences*, *93*(23), 12658-12663.

Jaffe, A. B., Fogarty, M. S., & Banks, B. A. (1997). *Evidence from patents and patent citations on the impact of NASA and other federal labs on commercial innovation* (No. w6044). National Bureau of Economic Research.

Jaffe, A. B., Newell, R. G., & Stavins, R. N. (2004). Technology policy for energy and the environment. *Innovation policy and the economy*, *4*, 35-68.

Jaffe, A. B., Newell, R. G., & Stavins, R. N. (2005). A tale of two market failures: Technology and environmental policy. *Ecological economics*, *54*(2), 164-174.

Jaffe, A. B., & Palmer, K. (1997). Environmental regulation and innovation: a panel data study. *Review of economics and statistics*, *79*(4), 610-619.

Jaffe, A., & Stavins, R.N. (1995). Dynamic incentives of environmental regulations: The effects of alternative policy instruments on technology diffusion. *Journal of Environmental Economics and Management, 29,* S-43-S-63.

Jaffe, A. B., & Trajtenberg, M. (1996). Flows of knowledge from universities and federal laboratories: Modeling the flow of patent citations over time and across institutional and geographic boundaries. *proceedings of the National Academy of Sciences*, *93*(23), 12671-12677.

Jaffe, A. B., & Trajtenberg, M. (2002). *Patents, citations, and innovations: A window on the knowledge economy*. MIT press.

Johnstone, N., Haščič, I., & Popp, D. (2010). Renewable energy policies and technological innovation: evidence based on patent counts. *Environmental and resource economics*, *45*(1), 133-155.

Joint Economic Committee of the United States Congress. (2010). *The Pivotal Role of Government Investment in Basic Research, A report to the United States Senate*.

Katila, R. (2000). Using patent data to measure innovation performance. *International Journal of Business Performance Management*, *2*(1-3), 180-193.

Katila, R. (2002). New product search over time: past ideas in their prime?. *Academy of Management Journal*, *45*(5), 995-1010.

Kemp, R. (1997). *Environmental policy and technical change: A comparison of the technological impact of policy instruments.* Cheltenham: Edward Elgar.

Kemp, R., & Pontoglio, S. (2008). The innovation effects of environmental policy instruments – a typical case of the blind men and the elephant. *Paper for DIME WP 2.5 Workshop on Empirical Analyses of Environmental Innovations, Fraunhofer Institute for Systems and Innovation Research (ISI), Karlsruhe*, January, 17th-18th.

King, G. (1986). How not to lie with statistics: Avoiding common mistakes in quantitative political science. *American Journal of Political Science*, 666-687.

King, G. (1988). Statistical models for political science event counts: Bias in conventional procedures and evidence for the exponential Poisson regression model. *American Journal of Political Science*, 838-863.

King, G. (1989a). Event count models for international relations: Generalizations and applications. *International Studies Quarterly*, *33*(2), 123-147.

King, G. (1989b). Variance specification in event count models: From restrictive assumptions to a generalized estimator. *American Journal of Political Science*, 762-784.

Klette, T., & Griliches, Z. (2000). Empirical patterns of firm growth and R&D investment: a quality ladder model interpretation. *The Economic Journal, 110 (463),* 363-387.

Klette, T., Moen, J., & Griliches, Z. (2000). Do subsidies to commercial R&D reduce market failures? Microeconomic evaluation studies. *Research policy, 29,* 471-495.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science, 3 (3),* 383-397.

Lanjouw, J. O., & Schankerman, M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators. *The Economic Journal*, *114*(495), 441-465.

Lee, J. (2006). Ph.D. Dissertation, Carnegie Mellon University, Pittsburg, PA

Lehmann, S., Jackson, A. D., & Lautrup, B. (2008). A quantitative analysis of indicators of scientific performance. *Scientometrics, 76 (2),* 369-390.

Leonard, T. C. (2009). "Redeemed by History" Review essay on Thomas K. McCraw, *Prophet of Innovation: Joseph Schumpeter and Creative Destruction*, Cambridge: Harvard University Press. *History of Economic Ideas XVII* (1)

Lerner, J., & Seru, A. (2015). The use and misuse of patent data: Issues for corporate finance and beyond. *Booth/Harvard Business School Working Paper*.

Libanio, G. A. (2005). Unit roots in macroeconomic time series: theory, implications and evidence. *Nova Economia,* 15 (3).

Long, C. (2002). Patent signals. *The University of Chicago Law Review*, 625-679.

Mansfield, E. (1986). Patents and innovation: An empirical study, *Management Science, 32(2),* 173-81.

Margolis, R., & Kammen, D. (1999). Evidence of under-investment in energy R&D in the United States and the impact of Federal policy. *Energy Policy*, 27 (10), 575-584.

Marinova, D., & McAleer, M. (2003). Modelling trends and volatility in ecological patents in the USA, *Environmental Modelling & Software*, 18 (3): 195-203

Meyer, M. (2001). Patent citation analysis in a novel field of technology: An exploration of nano-science and nano-technology. *Scientometrics,* 51 (1): 163-83

Mitchell, C., Bauknecht, D., & Connor, P.M. (2006). Effectiveness through risk reduction: a comparison of the renewable obligation in England and Wales and the feed-in system in Germany, *Energy Policy*, 34 (3), 297-305

Mowery, D.C., & Rosenberg, N. (1982). The influence of market demand upon innovation: A critical review of some recent empirical studies. In N. Rosenberg (Ed.), *Inside the Black Box: Technology and Economics* (pp. 193-241). New York: Cambridge University Press.

Murphy, K., Shleifer, A., & Vishny R. (1991). The Allocation of Talent: Implications for Growth, *The Quarterly Journal of Economics*, 106 (2): 503-530.

Nelson, C. R., & Plosser, C. R. (1982). Trends and random walks in macroeconmic time series: some evidence and implications. *Journal of monetary economics*, *10*(2), 139-162.

Nelson, R., & Winter, S. (1977). In search of a useful theory of innovation. *Research Policy, 6,* 36-76.

Nemet, G., & Kammen, D. (2007). U.S. energy research and development: Declining investment, increasing need, and the feasibility of expansion. *Energy Policy*, 35 (1), 746-755

O'Donoghue, T., Scotchmer, S., & Thisse, JF. (1998). Patent breadth, patent life, and the pace of technological progress. *Journal of Economics & Management Strategy, 7 (1),* 1–32.

OECD (1995). The Measurement of Scientific and Technological Activities. Proposed Guidelines for Collecting and Interpreting Technological Innovation Data. Oslo Manual. 2nd edition, DSTI, OECD / *European Commission Eurostat*, Paris.

Phillips, P. C., & Moon, H. R. (2000). Nonstationary panel data analysis: an overview of some recent developments. *Econometric Reviews*, *19*(3), 263-286.

Popp, D. (2002). Induced Innovation and Energy Prices. *American Economic Review*, 92 (1): 160-180.

Popp, D. (2005). Lessons from patents: Using patents to measure technological change in environmental models, *Ecological Economics*, 54 (2–3), 209-226

Popp, D. (2006). They don't make them like they used to: an examination of energy patent citations over time, The Economics of Innovation and New Technology, 15 (8), 753–776

Popp, D., R. Newell, & A. Jaffe. (2009). Energy, the environment, and technological change. *NBER Working Paper* 14832.

Raffalovich, L. E. (1994). Detrending time series A cautionary note. *Sociological Methods & Research*, *22*(4), 492-519.

Raytheon Service Corporation. (1980). *Wind Energy Systems Program Summary,* DOE/CS/20097-01 (Prepared for the USDOE).

Rogers, E.M. (1995). *Diffusion of innovations (*4th ed.). New York: Free Press.

Rosenberg, N. (1969). The direction of technological change: Inducement mechanisms and focusing devices. *Economic Development and Cultural Change, 18,* 1-24.

Sargan, J. D., & Bhargava, A. (1983). Maximum likelihood estimation of regression models with first order moving average errors when the root lies on the unit circle. *Econometrica: Journal of the Econometric Society*, 799-820.

Sargan, J. D., & Bhargava, A. (1983). Testing residuals from least squares regression for being generated by the Gaussian random walk. *Econometrica: Journal of the Econometric Society*, 153-174.

Sawin, J. (2001). *The Role of Government in the Development and Diffusion of Renewable Energy Technologies: Wind Power in the United States, California, Denmark, and Germany,*

1970-2000 (Ph.D. Thesis, Tufts University, Medford, MA). Received from Janet Sawin via email, January 13, 2017.

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy*. New York: Harper Brothers.

Scotchmer, S. (2004a). The political economy of intellectual property treaties. *The Journal of Law, Economics & Organization, 20 (2),* 415-437.

Scotchmer, S. (2004b). *Innovation and Incentives.* Cambridge: The MIT Press.

Smith, K. (1992a). *Quantitative innovation studies in Europe with existing datasets: Possibilities and problems*. Oslo: Royal Norwegian Council for Scientific and Industrial Research.

Smith, K. (1992b). Technological innovation indicators: Experience and prospects. *Science and Public Policy, 19 (6),* 383-92.

Smith, R. P. (1999). Unit roots and all that: the impact of time-series methods on macroeconomics. *Journal of Economic methodology*, *6*(2), 239-258.

Solar Energy Research, Development and Demonstration Act of 1974. (SERDDA). (1974). Public Law 473, 93rd Congress, second session.

Solar Energy Research Institute (SERI). (1987). *Wind Energy Systems: Program Summary Fiscal Years 1986 and 1987,* as cited by Sawin 2001.

Stephan, P., Gurmu, S, Sumell, A.J., & Black, G. (2002). Individual patenting and publication activity: having one's cake and eating it too. *Prepared for presenting at the NPRNet Conference 2002, Brighton,* March 23rd.

Stoneman, P. (Ed.). (1995). *Handbook of the economics of innovation and technological change.* Oxford, UK: Blackwell.

Stuart, T. E. (2000). Interorganizational alliances and the performance of firms: A study of growth and innovation rates in a high-technology industry. *Strategic management journal*, 791-811.

Swann, P. (1993). Inference from Mixed Bags: The Economic Value of Patent Counts, Innovation Counts and the Like. London Business School, working paper.

Taylor, M. (2008a). Beyond technology-push and demand-pull: Lessons from California's solar policy. *Energy Economics*, *30*(6), 2829-2854.

Taylor, M. R. (2008b). *Induced technical innovation under cap-and-trade abatement programs*. Working Paper, Goldman School of Public Policy. University of California, Berkeley.

Taylor, M., Nemet, G., Colvin, M., Begley, L., Wadia, C., & Dillavou, T. (2007). Government actions and innovation in clean energy technologies: The cases of photovoltaic cells, solar thermal electric power, and solar water heating. *California Energy Commission, California*.

Taylor, M.R., Rubin, E.S., & Hounshell, D.A. (2003). The Effect of Government Actions on Technological Innovation for SO2 Control. *Environmental Science and Technology*, 37 (20), 4527-4534.

Taylor, M.R., Rubin, E.S., & Hounshell, D.A. (2005). "Regulation as the Mother of Innovation: The Case of SO2 Control." *Law & Policy*, *27 (2),* 348-78.

Taylor, M.R., Thornton, D., Nemet, G., & Colvin, M. (2006). Government Actions and Innovation in Environmental Technology for Power Production: The Cases of Selective Catalytic Reduction and Wind Power in California. *Report to the California Energy Commission Public Interest Energy Research Program*. CEC-500-2006-053.

Tushman, M. L., & Anderson, P. (1986). Technological discontinuities and organizational environments. *Administrative science quarterly*, 439-465.

Tushman, M. & Murmann, J. P. (2002). Dominant Designs, Technology Cycles and Organizational Outcomes. In Garud, R., Kumaraswamy, A. and Langlois, R. N. (Ed.).*Managing in the Modular Age: Architectures, Networks, and Organizations*.Malden, MA, Blackwell.

Tushman, M. L., & Nelson, R. R. (1990). Introduction: Technology, organizations, and innovation. *Administrative Science Quarterly*, *35*(1), 1-8.

Tushman, M. L. & Romanelli, E. (1985). Organizational Evolution: A Metamorphosis Model of Convergence and Reorientation. In Cummings, L. L. and Staw, B. M. (Ed.).*Research in Organizational Behavior*.(Vol. 7).Greenwich, CT, JAI Press Inc.

Tushman, M. L., & Rosenkopf, L. (1992). Organizational determinants of technological-change-toward a sociology of technological evolution. *Research in organizational behavior*, *14*, 311-347.

United States Department of Energy. (USDOE). (1980). *Wind Energy Systems Program.*

West, K. (1988). On the interpretation of near random-walk behavior in GDP. *American Economic Review*, v. 78, n. 1, p. 202-209.

Williamson, O. (1979). Transaction Cost Economics; The Governance of Contractual Relations. *Journal of Law and Economics,* 22: 233-262.

Wind Energy Systems Act of 1980. (WESA). (1980). Public Law 345, 96th Congress, second session.

Yeh, S., Rubin, E.S., Hounshell, D.A., & Taylor, M.R. (2005). Technology Innovations and Experience Curves for NOx Control Technologies. *Journal of the Air & Waste Management Association, 55*, 1827–1838.

Yeh, S., Rubin, E.S., Hounshell, D.A., & Taylor, M.R. (2007). Uncertainties in Technology Experience Curves for Integrated Assessment Models. Accepted at *International Journal of Energy Technology and Policy*.

Zundel (Eds.). (2005). *Time Strategies, Innovation and Environmental Policy* (pp. 10-50), Cheltenham: Edward Elgar.

Zundel, S., Erdmann, G., Kemp, R., Nill, J., & Sartorius, C. (2005). Conceptual framework. *Sartorius, Christian/Zundel, Stefan (2005): Time Strategies, Innovation and Environmental Policy, Cheltenham*, 10-54.

Appendix A: Stata Code and Notes for Chapter 3

*Stationary Random Data for Graph, with Phillips-Perron Test and OLS Regression
clear
set seed 123456

quietly set obs 35
quietly gen t=_n
quietly tsset t
quietly gen K=rnormal(1000, 200)
pperron K, trend
reg K t, r

*Non-Stationary Random Data for Graph, with Phillips-Perron Test and OLS Regression
clear
set seed 12345678

quietly set obs 35
quietly gen t=_n
quietly tsset t
quietly gen y=rnormal(60, 30) if _n==1
quietly replace y=y[_n-1]+rnormal(-0.68, 12.86) if missing(y)
quietly replace y=0 if y<0
quietly replace y=round(y)
pperron y, trend
reg y t, r

*Tests of Non-Stationarity of Macro-Economic Data for Table 3.2.1
clear
import excel "testdata.xls"
quietly set obs 35
quietly gen t=_n
quietly tsset t
pperron A, trend
dfuller A, trend
pperron B, trend
dfuller B, trend
pperron C, trend
dfuller C, trend

*Tests of Non-Stationarity of Renewable-Energy Patent Data for Table 3.2.2
clear
use "R&D Paper Data.dta"
pperron Wnd_CC, trend
dfuller Wnd_CC, trend
pperron PV_CC, trend
dfuller PV_CC, trend
pperron STE_CC, trend
dfuller STE_CC, trend
pperron SWH_CC, trend
dfuller SWH_CC, trend

*Statistical Properties of Renewable-Energy Patent Data
clear
use "R&D Paper Data.dta"
sum Wnd_CC
sum PV_CC
sum STE_CC
sum SWH_CC

*OLS, Negative Binomial, and Log-Log Regressions for Renewable-Energy Patent Data and Macro-Economic Variables for Table
clear
use "R&D Paper Data.dta"
reg Wnd_CC GDP_2005D, r
reg PV_CC GDP_2005D, r
reg STE_CC GDP_2005D, r
reg SWH_CC GDP_2005D, r
reg Wnd_CC CWEFP_2000D, r
reg PV_CC CWEFP_2000D, r
reg STE_CC CWEFP_2000D, r
reg SWH_CC CWEFP_2000D, r
nbreg Wnd_CC GDP_2005D, r
nbreg PV_CC GDP_2005D, r
nbreg STE_CC GDP_2005D, r
nbreg SWH_CC GDP_2005D, r
nbreg Wnd_CC CWEFP_2000D, r
nbreg PV_CC CWEFP_2000D, r
nbreg STE_CC CWEFP_2000D, r
nbreg SWH_CC CWEFP_2000D, r
reg lnWndCC lnGDP, r
reg lnPVCC lnGDP, r
reg lnSTECC lnGDP, r
reg lnSWHCC lnGDP, r

73

reg lnWndCC lnCWEFP, r
reg lnPVCC lnCWEFP, r
reg lnSTECC lnCWEFP, r
reg lnSWHCC lnCWEFP, r

*Synthetic Data Means, Standard Deviations, Minimums, and Maximums

```
clear
timer on 1
set seed 52245

postfile buffer Mean SD Min Max using mcs, replace

forvalues i=1/10000 {
quietly drop _all
quietly set obs 35
quietly gen t=_n
quietly tsset t
quietly gen y=rnormal(60, 30) if _n==1
quietly replace y=y[_n-1]+rnormal(-0.68, 12.86) if missing(y)
quietly replace y=0 if y<0
quietly replace y=round(y)
quietly count if y<0.1
quietly sum y
quietly gen SD = r(sd)
quietly gen Mean = r(mean)
quietly gen Min = r(min)
quietly gen Max = r(max)
post buffer (Mean) (SD) (Min) (Max)
}

postclose buffer

use mcs, clear

timer off 1
summarize
timer list 1
timer clear 1
```

* OLS Regression on Levels Data Monte Carlo Simulation Code
```
clear
timer on 1
set seed 123456

postfile buffer tval using mcs, replace

forvalues i=1/1000 {
quietly drop _all
import excel "testdata.xls" /*comment: This file contains GDP and Energy Price data found
in Appendix B*/
quietly set obs 35
quietly gen t=_n
quietly tsset t
quietly gen K=rnormal(1000, 200)
quietly gen y=rnormal(60, 30) if _n==1
quietly replace y=y[_n-1]+rnormal(-0.68, 12.86) if missing(y)
quietly replace y=0 if y<0
quietly replace y=round(y)
quietly reg y A /*comment: use 'A' in this line of code and the next for GDP, 'B' for
Consumption Weighted Electricity Price Data, and 'K' substituted for 'y' for random
stationary data*/
quietly gen tval=_b[A]/_se[A]
post buffer (tval)
}

postclose buffer

use mcs, clear

timer off 1
summarize
count if tval<-1.96
count if tval>1.96
timer list 1
timer clear 1
```

76

* Log-Log Regression on Levels Data Monte Carlo Simulation Code
```
clear
timer on 1
set seed 123456

postfile buffer tval using mcs, replace


forvalues i=1/10000 {
quietly drop _all
import excel "testdata.xls" /*comment: This file contains GDP and Energy Price data*/
quietly set obs 35
quietly gen t=_n
quietly tsset t
quietly gen K=rnormal(1000, 200)
quietly gen y=rnormal(60, 30) if _n==1
quietly replace y=y[_n-1]+rnormal(-0.68, 12.86) if missing(y)
quietly replace y=1 if y<1
quietly replace y=round(y)
quietly gen lny = ln(y)
quietly gen lnA = ln(A)
quietly gen lnB = ln(B)
quietly gen lnK = ln(K)
quietly reg lny lnA /*comment: use 'lnA' in this line and the next for GDP, 'lnB' for
Consumption Weighted Electricity Price Data, and 'lnK' instead of 'lny' for random
stationary data*/
quietly gen tval=_b[lnA]/_se[lnA]
post buffer (tval)
}

postclose buffer

use mcs, clear

timer off 1
summarize
count if tval<-1.96
count if tval>1.96
timer list 1
timer clear 1
```

* Negative Binomial Regression on Levels Data Monte Carlo Simulation Code
```
clear
timer on 1
set seed 123456

postfile buffer tval using mcs, replace

forvalues i=1/10000 {
quietly drop _all
import excel "testdata.xls" /*comment: This file contains GDP and Energy Price data found
in Appendix B*/
quietly set obs 35
quietly gen t=_n
quietly tsset t
quietly gen K=rnormal(1000, 200)
quietly gen y=rnormal(60, 30) if _n==1
quietly replace y=y[_n-1]+rnormal(-0.68, 12.86) if missing(y)
quietly replace y=0 if y<0
quietly replace y=round(y)
quietly nbreg y A /*comment: use 'A' in this line and the next for GDP, 'B' for Consumption
Weighted Electricity Price Data, and 'K' substituted for "y" for random stationary data*/
quietly gen tval=_b[A]/_se[A]
post buffer (tval)
}

postclose buffer

use mcs, clear

timer off 1
summarize
count if tval<-1.96
count if tval>1.96
timer list 1
timer clear 1
```

* OLS Regression on Changes Data Monte Carlo Simulation Code
```
clear
timer on 1
set seed 12345

postfile buffer tval using mcs, replace

forvalues i=1/10000 {
quietly drop _all
import excel "testdata.xls" /*comment: This file contains GDP and Energy Price data found
in Appendix B*/
quietly set obs 35
quietly gen t=_n
quietly tsset t
quietly gen y=. if _n==1
quietly replace y=rnormal(-.68,12.86) if missing(y)
quietly replace y=. if _n==1
quietly replace y=round(y)
quietly reg y d.A, r /*comment: use 'A' in this line and the next for GDP, 'B' for
Consumption Weighted Electricity Price Data*/
quietly gen tval=_b[d.A]/_se[d.A]
post buffer (tval)
}

postclose buffer

use mcs, clear

timer off 1
summarize
count if tval<-1.96
count if tval>1.96
timer list 1
timer clear 1
```

*Unit-Root Test on Synthetic Data – Levels and Changes
clear
timer on 1
set seed 123456

postfile buffer pval using mcs, replace

forvalues i=1/10000 {
quietly drop _all
quietly set obs 35
quietly gen t=_n
quietly tsset t
quietly gen y=rnormal(60, 30) if _n==1
quietly replace y=y[_n-1]+rnormal(-0.68, 12.86) if missing(y)
quietly replace y=0 if y<0
quietly replace y=round(y)
quietly pperron d.y, trend **/*comment: this code runs Phillips Perron test on changes data - remove the 'd.' in this line to run on levels data*/**
quietly gen pval=r(pval)
post buffer (pval)
}

postclose buffer

use mcs, clear

timer off 1
summarize
count if pval<0.05
timer list 1
timer clear 1

Appendix B: Data for Models and Simulations in Chapter 3

The first column is GDP, the second is the consumption weighted electricity fuel price data, and the third is the S&P 500 index for 1973 to 2008. All three series are inflation adjusted as described in chapter 3.

GDP	Electricity fuel price	S&P 500
4889.9	2196.889	356
4879.5	2439.051	349
5141.3	2464.743	389
5377.7	2484.708	346
5677.6	2532.353	315
5855	2743.216	306
5839	3053.842	312
5987.2	3276.768	303
5870.9	2994.444	268
6136.2	2737.478	347
6577.1	2555.369	332
6849.3	2334.119	378
7086.5	1917.724	469
7313.3	1807.182	542
7613.9	1663.812	488
7885.9	1663.626	567
8033.9	1623.696	548
8015.1	1492.751	603
8287.1	1432.219	640
8523.4	1386.533	676
8870.7	1304.445	669
9093.7	1194.03	773
9433.9	1276.347	926
9854.3	1271.367	1175
10283.5	1161.617	1437
10779.8	1191.432	1720
11226	1533.928	1775
11347.2	1586.005	1442
11553	1365.841	1183
11840.7	1693.864	1133
12263.8	1902.641	1292
12638.4	2418.021	1331
12976.2	2211.508	1408
13254.1	2262.738	1535
13312.2	2730.75	1215

Appendix C: Methods Described in Chapter 3 Applied to Panel Data

Some researchers have claimed that panel data avoids spurious regressions (i.e. elevated false positives); however, while this is true in the asymptote (Phillips and Moon 2007), I see no reason to believe it is accurate in the relatively small panels frequently used in economic and organizational research. In an attempt to generalize the time-series findings from chapter 3, I ran a modified version of my Monte Carlo simulations (see the sample code below), designed to generate panel data composed of a variable number of random walks. For speed of analysis I limited myself to a panel with 35 years of data (as with my previous simulations), five "entities" (in typical panel-data research these would be firms, countries, regions, or some other well defined unit of analysis over which one could count outcomes of interest), and 1,000 simulations.

Findings are shown in Table C.1 below.[6] OLS and negative binomial regressions both yielded high rates of spurious correlations. The number of entities compiled in the panel can be increased and will still generate similar results – I tested this with small numbers of runs, but did not feel compelled to do so in a thorough way, and so do not present the results here. An expanded version of this work may be worthwhile, as this demonstrates my findings from chapter 3 generalize to panel data.

Table C1: False-Positive* Rates for Synthetic Panel Data and Macro-Economic Variables

	OLS		Neg. Binomial	
(n=1000)	USGDP	Energy Prices	USGDP	Energy Prices
Non-Stationary Synthetic Panel Data	67.8%	56.5%	72.3%	56.8%
Changes on Non-Stat. Panel Data	4.9%	6.1%		

*False positives here are P values less than 0.05 – the expected rate of false positives is 5%.

6

Monte Carlo simulations for the negative binomial regressions required segmentation and use of multiple seeds for STATA's random number generator. Large numbers of simulated regressions (typically upwards of 100 in the case of panel data runs) tended to eventually generate one regression that would become "trapped" in a convexity or discontinuity during successive iterations and thereby foul the entire run. Segmentation allowed for smaller runs that avoided fouling, from which I then summed the results for a full total of 1,000 runs.

Sample Stata Code for Monte Carlo Simulations of Panel Data

```
clear
timer on 1
set seed 12345678

postfile buffer tval using mcs, replace


forvalues i=1/1000 {
quietly drop _all
import excel "testdata.xls"
quietly set obs 35
quietly gen t=_n
forvalues j=1/5 {
quietly gen y`j'=rnormal(60, 30) if _n==1
quietly replace y`j'=y`j'[_n-1]+rnormal(-0.68, 12.86) if missing(y`j')
quietly replace y`j'=1 if y`j'<1
quietly replace y`j'=round(y`j')
}
quietly reshape long y, i(t) j(id)
quietly xtset id t
quietly xtreg y A
quietly gen tval=_b[A]/_se[A]
post buffer (tval)
}

postclose buffer

use mcs, clear

timer off 1
summarize
count if tval<-1.96
count if tval>1.96
timer list 1
timer clear 1
```
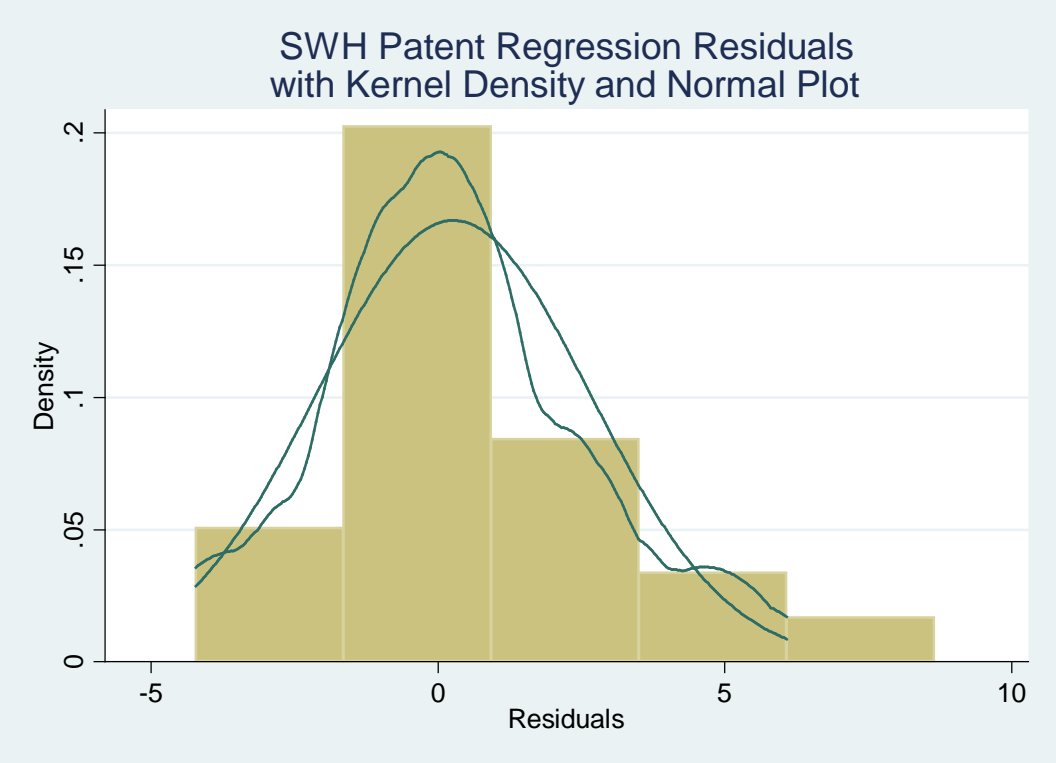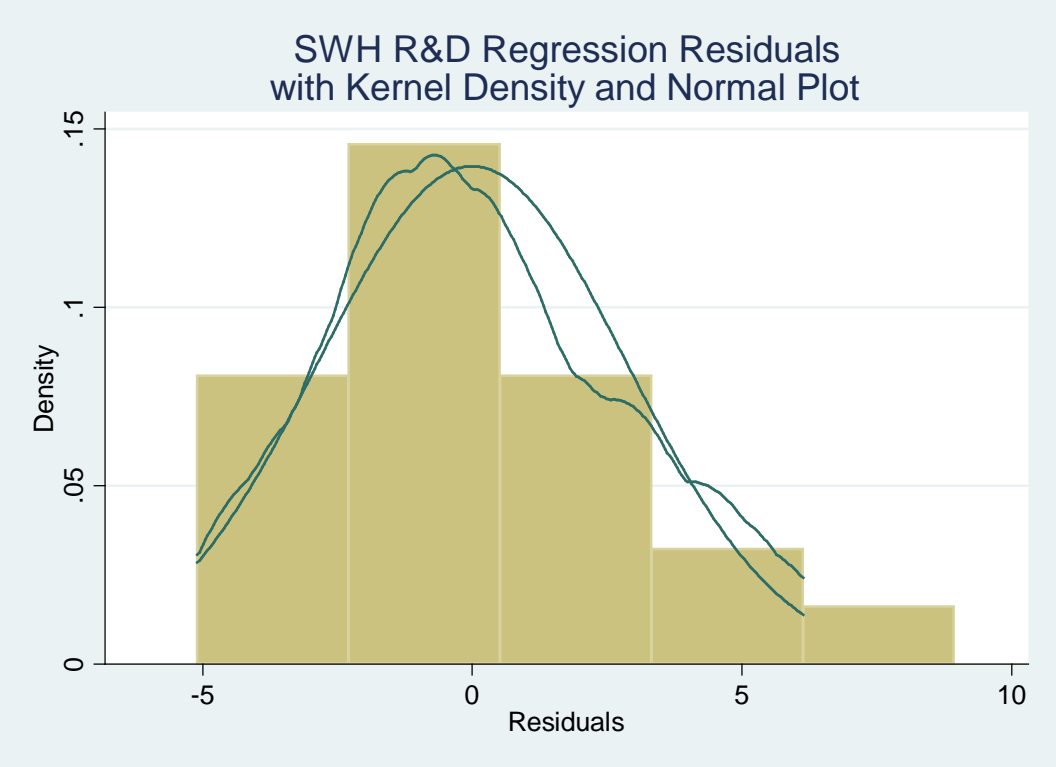
Appendix D: Residuals Histograms for Selected Renewable Energy VAR Models in Chapter 4
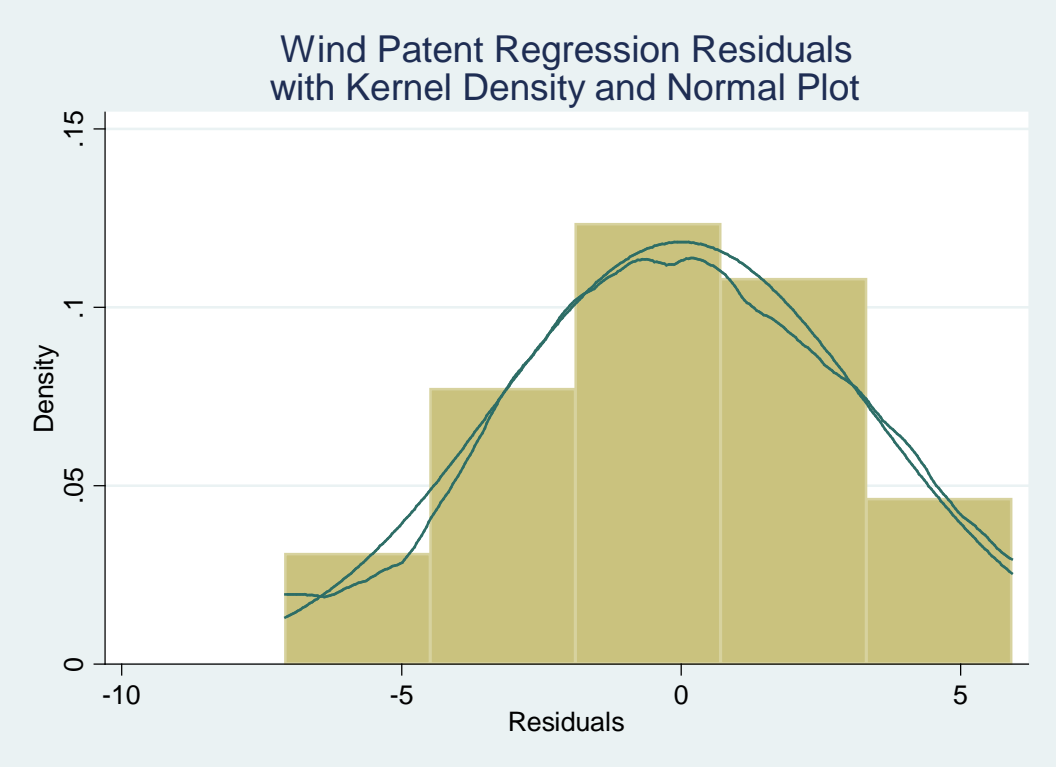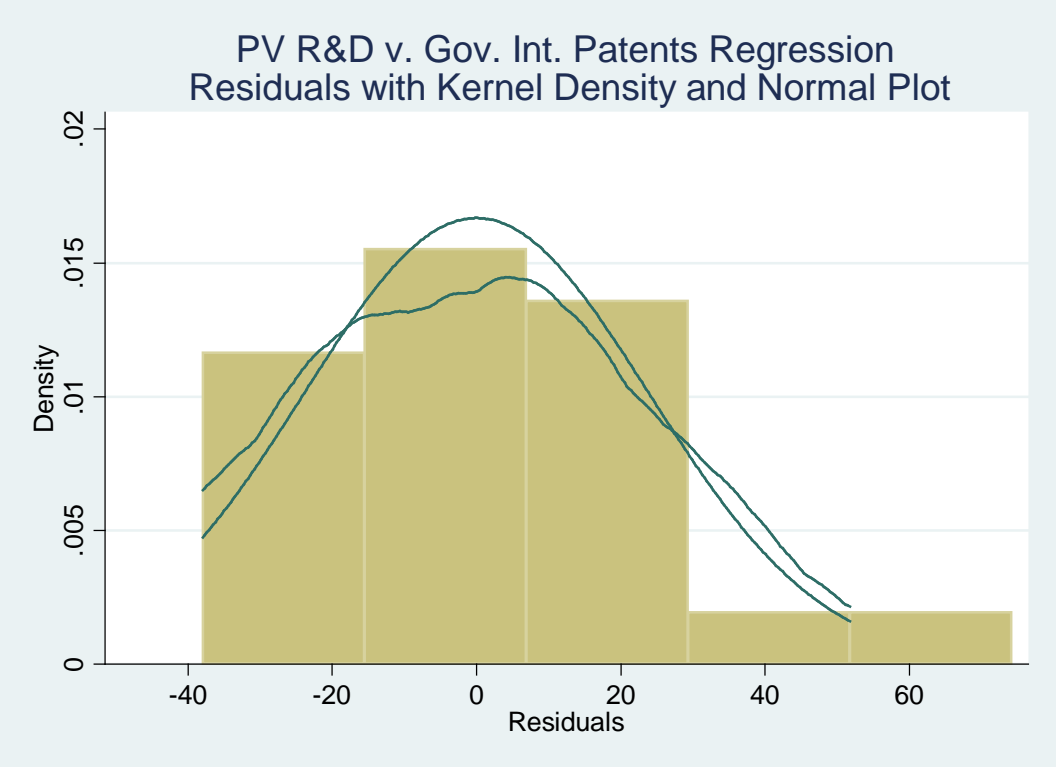
Graph D1:

SWH Patent Regression Residuals with Kernel Density and Normal Plot

Graph D2:

SWH R&D Regression Residuals
with Kernel Density and Normal Plot

Graph D3:

Wind Patent Regression Residuals
with Kernel Density and Normal Plot

Graph D4:

PV R&D v. Gov. Int. Patents Regression
Residuals with Kernel Density and Normal Plot

Graph D5:

STE Gov. Int. Patents v. R&D Regression
Residuals with Kernel Density and Normal Plot