Title
Assessing Low-Carbon Fuel Technology Innovation Through a Technology Innovation System Approach

Author
Kessler, Jeff

Publication Date
2015-12-01

Research Report – UCD-ITS-RR-15-26

Assessing Low-Carbon Fuel Technology Innovation Through a Technology Innovation System Approach

December 2015

Jeff Kessler

Institute of Transportation Studies ◦ University of California, Davis
1605 Tilia Street ◦ Davis, California 95616
PHONE (530) 752-6548 ◦ FAX (530) 752-6572
www.its.ucdavis.edu

Assessing low-carbon fuel technology innovation through a Technology Innovation System approach

by

JEFF KESSLER
B.S. (University of Colorado) 2008
M.S. Chemical Engineering (University of Colorado) 2009

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTORS OF PHILOSOPHY

in Transportation Technology and Policy

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

____

Daniel Sperling, Chair

____

Joan Ogden

____

Sonia Yeh

Committee in Charge 2015

Abstract

Addressing anthropogenic climate change will require a variety of novel technology solutions. Where will these solutions come from, and how can we foster their development? To answer these questions, it is important to delve into the process of technology innovation. We need to better understand how technological transitions happen, and we need to figure out how innovation can be directed.

While the existing work on technology innovation is abundant, the innovation process largely remains a "black box," shrouded in mystery. Energy models that incorporate innovation concepts, such as experience curves, fail to consider the fundamental processes that drive innovation. More nuanced approaches to innovation, however, are largely qualitative and difficult to model or to employ. This makes it hard to draw objective conclusions, or to make predictions about technologies moving forward.

This dissertation research establishes a set of methodological approaches to better break in to this innovation black box, aiding in the quantification of the more qualitative approaches to innovation. These methods are applied to better examine low-carbon technology innovation in transportation. Specifically, this dissertation looks at biofuel innovation and the more recent diffusion of electric vehicles.

Patent trends, one traditional approach for quantifying innovations, are used to provide a point of comparison for the novel methodologies employed. This research shows that the innovation narrative and conclusions that can be drawn from patent data are largely dependent on how patents are classified.

Employing statistical models in conjunction with computational linguistics and machine-learning algorithms, it is possible to classify large bodies of text. This

methodology is applied to a large selection of patents to better classify biofuel technologies. Additionally, this method is applied to a large repository of textual media, such as newspaper articles and trade journals, to select for specific technologies, and to classify articles by the type of information they convey. This Technology Innovation System (TIS) database is believed to adequately proxy the flow of information over time, due to the large number of documents collected.

The innovation trends captured in the TIS database align well with the biofuel narrative established in literature. There is also good alignment between patent data classified through this methodology and the TIS database.

Through use of the TIS database in conjunction with deployment data and policy data, this dissertation demonstrates several applications for assessing technology innovation. Results can be used to provide suggestions, supported by the data, which may foster improved innovation outcomes for low-carbon transportation technologies.

Acknowledgment

The path to writing and completing a dissertation is long, arduous, and at times downright frustrating. Without the support of a large network of individuals, I find it unlikely that I would have been able to sit still or remain focused long enough to follow this research project through from start to something that approximates "finished."

Even before enrolling at UC Davis, I received an unbelievable amount of support from Dan Sperling. I suspect that there are few academic advisors with his clout that are willing to meet at a random parking lot in the middle of nowhere to convince you to be their graduate student. Dan's mentorship, guidance, and openness to trying some world-class beer has been indispensible.

The Institute of Transportation studies and the myriad of excellent faculty and staff employed there have provided support at all hours in the morning and throughout this process. Specifically I would like to acknowledge the tremendous support I've received from Gustavo Collantes, Sonia Yeh, Joan Ogden, and Annemarie Schaaf.

I have been unbelievably fortunate to be fully funded and supported during my time working on this dissertation. The majority of funding support for my dissertation is from the National Science Foundation, specifically the GK-12 program through UC Davis (Renewable Energy Systems Opportunity for Unified Research Collaboration and Education). This program directly allowed me to work with middle school classrooms in the greater Sacramento area creating a two-way science dialogue. Additionally, my dissertation was supported by a very generous fellowship from The University of California Transportation Center, and through the National Center for Sustainable Transportation.

The Davis community has provided me with friendship and support at the times I've needed it most. The Davis Short Story Club and my Farmhouse family have been the best group of people and friends anyone could ask for. One miss Alexandra Weill is also terrific for spending weeks with me in small, cold mountain tents – typical research obligations for most dissertation work.

I would also like to thank Zsuzsanna Csereklyei for her helpful comments and useful discussions on modeling. Last but not least, I would like to thank Dominique Meroux and Arthur Wu, two new students (at least at the time of this writing) in the Transportation Technology and Policy program. Both were willing to take the time to learn a little bit about the methodologies I employ in this dissertation, and to assist in training a computer to recognize the esoteric.

To all of you, I am indebted. Thanks!

Table of Contents

Table of Tables

Table of Figures

List of Abbreviations

AFDC – Alternative Fuel Data Center

AFV – Alternative Fuel Vehicle

BFTIS – Biofuel Technology Innovation System

DOE –Department of Energy

EIA – Energy Information Administration

EPA – Environmental Protection Agency

EV – Electric Vehicle

GHG – Greenhouse Gas

LCFS – Low Carbon Fuel Standard

NLP – Natural Language Processing

PHEV – Plug-in Hybrid Electric Vehicle

RFS – Renewable Fuel Standard

TIS – Technology Innovation System

Chapter 1: A Brief Discussion of Technology Innovation

This is an original work at the intersection of economics and engineering; what it lacks in one, it attempts to make up for in the other. In short: this is an interdisciplinary dissertation. The goal and aim of this work is to provide a resource and a new set of methods for researchers and policy makers that, when applied, may better inform their decisions regarding low-carbon technologies and transition strategies.

I am of the belief that the most promising route forward for addressing climate change is through the diffusion of low-carbon technologies that can compete with, and ultimately displace fossil fuel alternatives (e.g. Grubler, 2012; J. Jenkins et al., 2012; J. D. Jenkins, 2014; T. Nordhaus, Shellenberger, & Navin, 2008). For this to occur, we need to learn how to better direct and promote low-carbon technology innovation.

In this section I provide a brief discussion of innovation and how the study of innovation has evolved over time, predominately within the economics literature. I then briefly look at the methods commonly used for forecasting technology adoption and diffusion – engineering tools, in kind.

It is my intent to show that common economic and engineering approaches for incorporating innovation lack in explanatory or predictive power, and that new methods and approaches for assessing innovation are needed. Additional tools are also necessary to evaluate policy efficacy in promoting innovation, especially for targeted end-goals such as low-carbon technology transitions. Important to this discussion is an understanding of what technology innovation is, which I cover in Chapter 1.

To further explore this concept of technology innovation, I develop and build out my own methods for innovation assessment in Chapter 2. To provide a basis for assessment, I examine a traditional approach for evaluating technology innovation, patent activity and trends. My work shows that the innovation narratives supported by patent activity are largely dependent on how patents are classified and grouped together. I employ machine-learning algorithms in an attempt to improve on this weakness for the case of biofuels. Moving from patent activity, I continue to employ machine-learning methods to facilitate the use of a novel data set for assessing technology innovation, text-based news sources. In Chapter 4 I examine how well the innovation narrative is captured through textual media, and compare this to what patenting activity tells us. I further go on to explore the case of California, and how policies may impact the overall innovation system and technology outcomes.

In Chapter 5, I employ deployment data for electric vehicles alongside textual media and meetings with experts. Combined, these data help further vet the methodologies I put forward for assessing technology innovation. I find that there is reasonable alignment between the expert interviews and the innovation system approached that I employ. I further demonstrate the possibility of linking textual media to vehicle deployment data. This is one approach that my methodology enables which can help support or direct state-level policy or plans of action going forward for fostering electric vehicle adoption.

Without further ado, I turn to the question of innovation.

1.1 What is Innovation?

Innovation is often used to refer to the commercialization of new products or processes – those that mark a substantial departure from their known predecessors

(Fagerberg, 2003). The smart phone, for instance, is an icon of innovation; it marks not only a departure from traditional cell phones (which marked a departure from landline technology), but also a departure from personal computing devices (such as laptops). Although the smart phone is the embodiment, of innovation, it is not fully indicative of the innovation process that led up to its creation and eventual commercialization. Ultimately, the technology innovation process must be placed in a broader context.

Innovation allows for new things to happen, including social and economic shifts. Joseph Schumpeter's early research into economic development led to his seminal works on innovation (Schumpeter, 1934, 1942). Schumpeter established the importance of innovation in driving economic growth. Schumpeter also ingrained in society the role of the entrepreneur as a driver of innovation. At what point, however, does innovation occur, and when is something innovative?

Schumpeter's work established a very clear distinction between invention and innovation. "Invention" is framed as the development of an original idea for a new product or process, while "innovation" is the conversion of that product or process into something commercially viable that is subsequently diffused. In other words, for innovation to occur, both invention and the exploitation of that invention is required (Popp, 2005).

While intuitively it is easier to understand innovation as the development and use of physical products, that is not always the case. New processes or methods may also be innovative, and can contribute substantially to economic growth. Similarly, the establishment of governments, regulation, and law in and of itself may be innovations. Drawing from the work of Carlsson and Stankiewicz (1991), I choose to define innovation here as: the creation, dissemination, and use of new information. Physical systems and products can then be viewed as carriers or embodiments of innovation, as opposed to being "the" innovation itself.

I choose to frame innovation like this because it stresses the importance of information flow while deemphasizing the physical product. Revisiting the innovation example of the smart phone, it becomes apparent that the smartphone is a product of innovation – an instance where accumulated knowledge and information culminate in one specific use case. The success of the smartphone required the flow and exploitation of information across different areas, ranging from consumer acceptance and use of hand-held electronics to streamlined manufacturing and internet connectivity. In turn, the smartphone has contributed to the creation and dissemination of new information that can be used in new ways to further facilitate innovation.

It is important to recognize that under my framing of innovation, the innovation is not the smartphone itself, but the utilization and grouping together of knowledge and information to achieve a novel result. It is the exploitation and aggregation of information and knowledge associated with the smartphone that is innovative. From this framing, a natural transition to a network view of the innovation process is possible: nodes represent knowledge and information, and the structural components (like companies or actors) that contribute to the flow of knowledge connect nodes together. Figure 1.1 visualizes what a network mapping might look like for technologies that undergo incremental innovations, compared to those that undergo a creative destruction process, where a novel network emerges that ultimately displaces the previous technological system. Similarly, "hype" might be represented by a small network of knowledge nodes, where connectivity is substantial only for a few nodes, but not the larger network. This could ultimately lead to the perception that the network is larger and better connected than it is.

While the iPhone may be one physical embodiment of innovation, numerous other smartphone models are all derived from the same core information and knowledge flows. Once knowledge and information is sufficiently accumulated and connected within a network, the physical manifestation of this innovation is trivialized. If one network link were to drop out, it is likely that another actor would emerge to recreate that link, if the system is supported well enough.

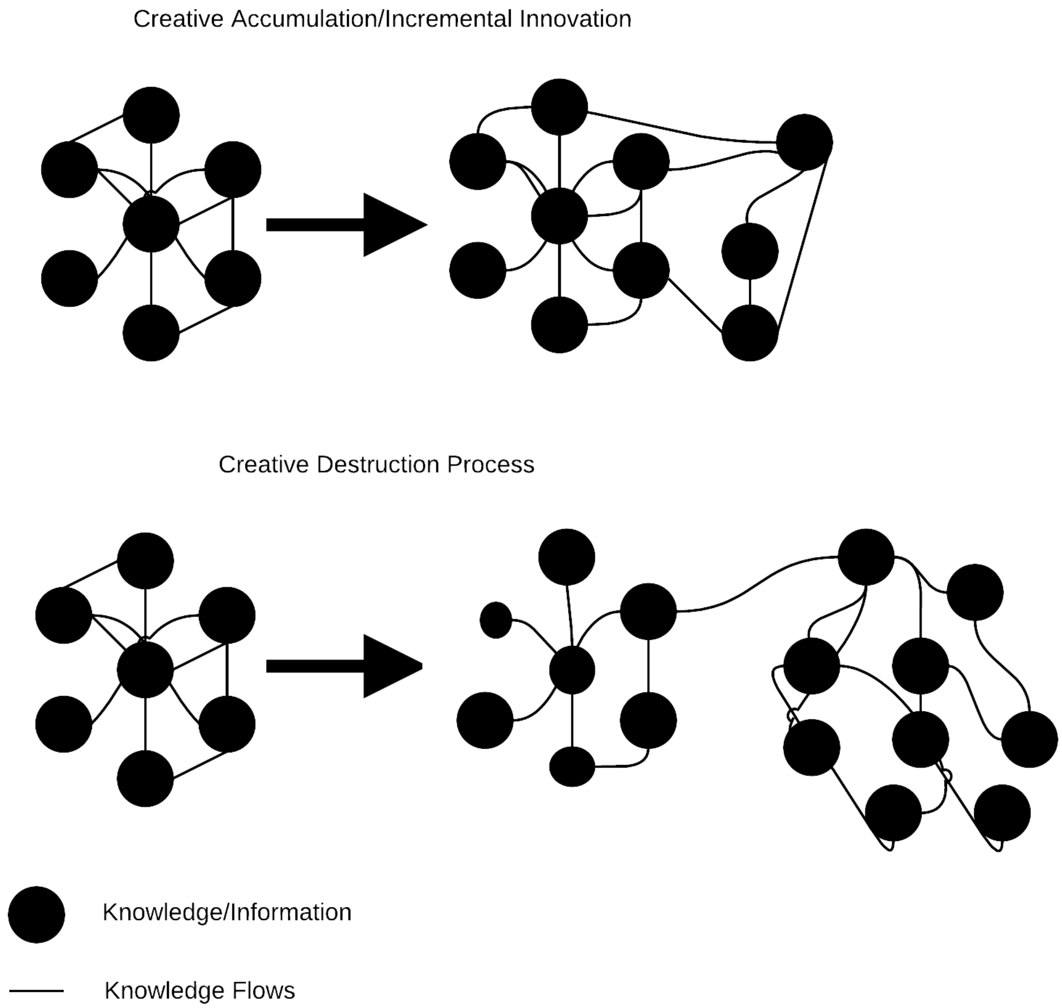

Figure 1.1 Stylized representation of an information network and the process that may ultimately contribute to incremental innovation, or to creative destruction.

1.2 The Study of Innovation

To better understand how innovation occurs, I first look at the body of research concerned with economic growth. This research motivated the study of innovation, leading to work aimed to establish the factors that drive and contribute to the innovation process.

Starting in 1956, Robert Solow released his pioneering work on economic growth, where he identified two explicit factors that drive economic growth: capital and labor

(Solow, 1956, 1957). Solow's novel contribution to growth economics, however, was the identification of a third, implicit factor: technical change.

While Solow's work considerably improved the model for economic growth, technical change was left as a "black box," with no formalization for how technical change manifests. It is also from this third, poorly understood factor that most economic growth stems (Solow, 1994). Due to this weakness, the Solow model has often been caricatured as treating technical change as "manna from heaven" (Audretsch, 2007). The Solow model recognizes that innovation drives technical change, thereby contributing to economic growth, but the factors that drive innovation are not captured.

Since Solow's work, there have been attempts to better model and endogenize the innovation "black box" (Paul M. Romer, 1986; Paul M Romer, 1990). Even after expansion of Solow's model, it still remains difficult to study the process of technical change, and the specific innovations that bring it about. Economics does not adequately capture the drivers of technology innovation, and is unable to answer the question "what happens after investment into R&D is made?" (Solow, 1994). This remains a weakness of neoclassical economics.

One way to remedy the weakness of the neoclassical models is to start innovation analysis at the micro level, taking into account the importance of the entrepreneur and the firm-specific decisions that drive innovation (Solow, 1994). The neoclassical models of innovation assume that firms have perfect knowledge and are able to perfectly optimize their behavior (Carlsson & Eliasson, 1995). Firms, however, do not operate with perfect knowledge, and they often utilize different knowledge bases, skillsets, and assumptions concerning technology and markets; this heterogeneity between firms merits a different theoretical framework for analyzing innovation (Carlsson & Stankiewicz, 1991). Historical case studies of innovative technologies can facilitate this process (Fouquet & Pearson, 2012; Grubler, 2012).

Ultimately, the need to better understand how innovation manifests has culminated in the development of a systems approach to thinking. This systems approach is more dynamic, and places technology at the forefront, building upon Schumpeter's second, and less-appreciated work on innovation that emphasizes the effects of actor-agent networks (Schumpeter, 1942). This view broadens the definition of the entrepreneur, allowing for a network of actors and agents to play a role in the innovation process.

At its core, the Schumpeterian view of innovation recognizes that there are two main patterns of innovation: a creative destruction pattern and a creative accumulation pattern (Stefano Breschi, Malerba, & Orsenigo, 2000; R. R. Nelson & Winter, 1982).

Creative destruction is a process in which established technologies, sectors, or firms are overturned ("destroyed") by newer, superior technology ("creation"). Given how I've defined innovation, I think it is worth conceptualizing this view in terms of information flow. Creative destruction embodies the process in which one network is destabilized and starts to diminish due to the creation of a newer network that supports or promotes a given technology. For instance, new information no longer promotes and connects to as many previous points of knowledge, and instead integrates into a different and distinct network, often at odds with the previous network of knowledge flows.

In contrast, creative accumulation is the process where sectors continue along the same trajectory, making incremental improvements in the technologies at hand. Or in a network view, the network continues to evolve and embody new actors and institutions, with new nodes and connections that are still strongly integrated into the existing network.

Schumpeter's work has fostered considerable study into innovation, leading to a systems view of the innovation process, in which the conventional entrepreneur is just one actor within a more complicated system. Carlsson and Stankiewicz (1991) define innovation systems as:

A network of agents interacting in a specific technology area under a particular institutional infrastructure to generate, diffuse, and utilize technology. Technological systems are defined in terms of knowledge or competence flows rather than flows of ordinary goods and services. They consist of dynamic knowledge and competence networks.

The innovation system approach recognizes that innovation is the product of an interactive process across a wide variety of actors. This process is not solely driven by the independent efforts of one entrepreneur. Stress is placed on the importance of a network, recognizing that firms do not innovate in isolation (Malerba, 2005). Institutions also shape firm innovation and interactions, which includes policies, regulations, and governments (Carlsson & Eliasson, 1995; Lundvall, Johnson, Andersen, & Dalum, 2002; R. Nelson, 1993).

The systems approach to innovation provides a new framing, in which the development of a network is fundamental. Networks act to bring about new technologies and to defend incumbent technologies. From this lens, it is possible to ascertain what developments take place that ultimately lead to innovation. A specific entrepreneur, product, or innovation is treated as the result of a strong network. To put this in the transportation context: with a sufficiently strong innovation network, if Ford had not employed the assembly line process for automobiles, it is inevitable that another entrepreneur would have successfully employed a similarly effective process.

To better understand the importance of networks in driving innovation, I discuss several different innovation systems frameworks and how they approach the innovation process.

1.3 The Innovation Process

Researchers have used case studies and historical accounts of technology development and diffusion to better study and understand the innovation process (e.g. Bergek & Jacobsson, 2003; Geels, 2002; Grübler, Nakićenović, & Victor, 1999). Case studies have helped researchers develop, and ultimately establish a set of commonalities and themes for technologies, firms, and sectors as part of the larger innovation process (Grubler, 2012; Johnson, 2001).

From these case studies, two dominant approaches to innovation assessment have emerged. Some researchers utilize historic relationships to explore deployment-focused, predictive models for technology diffusion, often relying on experience-curve relationships to improve neo-classical models (Grübler et al., 1999). Alternatively, researchers have also used case studies for more qualitative innovation assessment; an approach that promotes system-based frameworks and thinking. The qualitative innovation frameworks have largely been used to study innovation at national, regional, sectoral and technological levels (Johnson, 2001).

These qualitative frameworks for innovation help determine which details and features are important for achieving a given innovation objective – be it to promote growth within a specific sector or even a specific firm, or to promote technological transitions at a national level.

Predictive models for innovation often rely on experience curves. These models tend to treat the very factors that explain the experience curve as exogenous, relying on "typical" scenarios and assumptions about what might happen, rather than incorporating the factors that can influence or drive the "typical" values (Grübler et al., 1999). While informative, models that merely incorporate experience curves and learning effects are likely inadequate for exploring what can be done to change or accelerate the innovation process.

Because my research is primarily concerned with aiding the advancement of low-carbon technologies, as opposed to aiding advancement of all technology or promoting economic growth in general, I choose not to discuss some of the broader innovation system frameworks, such as the national or regional innovation systems (Lundvall, 1992, 2007; Lundvall et al., 2002). Instead, I limit my focus to the set of frameworks and analytical methods that deal with specific technology innovations.

What follows is a brief discussion of the most pertinent technology innovation system frameworks that I have come across in my research on the topic of innovation. I also summarize some of the case studies to which these frameworks have been applied. Table 1.1 serves to summarize the approaches to innovation that I look at here

Table 1.1

Summary of different technology innovation assessment frameworks and methods

	Sectoral Systems of Innovation	**Socio-Technical Systems**	**Technology Innovation Systems**	**Experience Curves**	**Diffusion Curves**
Summary	Qualitative assessment of specific firms in an already-existing network working to diffuse and develop a novel technology	Qualitative assessment of the user-groups that gradually adopt and utilize a novel technology	Qualitative assessment of the networks of actors and agents that exist and come into existence to utilize, diffuse, and develop a novel technology	Mathematical relationship that predicts how technology cost decreases based on the diffusion of that technology.	Mathematical relationship that determines the market penetration of a technology with respect to time.
Benefits	Allows for in-depth consideration of the current agents and actors working to promote a specific technology.	Allows for in-depth consideration of technology use cases. This framework enables niche-marketing strategies and thinking to facilitate technology diffusion	Allows for consideration of different networks of actors and agents that promote or block technology innovation. Actors and institutions are mapped to functions to better explore the innovation process	Can be incorporated into models to better approximate the process of innovation responsible for cost reductions	Can be incorporated into models to better approximate the rate at which a new technology can be expected to enter the market

Weaknesses	Does not consider blocking effects that hinder innovation outcomes, nor can it account for out-of-network effects that could promote or detract from successful innovations. The framework is largely qualitative.	Does not consider blocking effects that work against innovation, focusing instead on a demand-pull view to innovation. The framework is largely qualitative.	Reduces the innovation process down to a set of "functions," or things that have to exist for innovation to occur. Functional fulfillment is difficult to assess. The framework is largely qualitative.	The rate of cost reduction is based on historical accounts of similar technologies and expected outcomes. Because the relationship is logarithmic, small differences in curve expectations can result in order of magnitude discrepancies with actuality. The experience curve relationship does not endogenize all factors of the innovation process, and assumes that deployment alone is sufficient for reducing costs	The parameters that effect the rate of diffusion are highly variable, and approximated based on past technology diffusion events that may or may not accurately reflect the true rate of diffusion for the technology being modeled

There are three prevalent and influential technology innovation frameworks used in the literature to analyze and assess the innovation process for specific technologies (Coenen & Díaz López, 2010). These frameworks are: (1) Sectoral Systems of Innovation (SSI), (2) Social Technical Systems (ST-Systems) of innovation, and (3) Technology Innovation Systems (TIS). Each of these frameworks has its advantages and disadvantages, and each framework can be applied in its own right to energy transitions and the adoption of low-carbon technologies.

The SSI framework, for instance, looks at how existing firms operate in a space to promote or adopt new technologies. For instance, the SSI framework could be applied to better reveal for how existing, large automotive manufacturers, like Ford or General Motors, contribute to electric vehicle adoption through the incremental adoption of electrification in drivetrains. For instance, increased fuel efficiency standards can lead to increased vehicle hybridization, which eventually leads to fully electric vehicle models.

The ST-System of innovation instead looks at how technologies diffuse through society, moving from one user-group to another, slowly improving as new user-groups find novel applications for the technology. For instance, this framework can be used to better understand how second generation biofuel innovation is being promoted through niche-adoption and use in the military. The military may find that biofuels offer considerable supply chain advantages for combat situations compared to fossil fuels. In turn, learning and scaling up of the technology for these military operations may find further application in the aviation sector, before being deployed at a much larger commercial scale.

The TIS framework, in comparison, looks at how a network of actors and institutions develops over time to use, improve on, and further diffuse a technology. The

TIS approach helps to reveal where weaknesses exist in the network that can block or hinder innovation outcomes. TIS incorporates elements from both the ST-Systems and SSI frameworks. For example, the TIS framework could be used to discuss the emergence of a set of actors and policies, and how they interact with one another and in the context of the larger network to promote or detract from electric vehicle innovation. This could include assessment of manufactures, like Tesla, and policies like the Zero Emission Vehicle Mandate in California. In turn, the TIS framework can be used to facilitate discussion about how an incumbent network, such as large automotive manufactures, are hindering or contributing to the development or deployment of this new network (e.g. lobbying against policies, or hindering vehicle adoption through existing dealership models).

I further explore the use of and background for each of these frameworks below.

1.3.1 Sectoral Systems of Innovation

The SSI framework places system dynamics and process transformation at the center of analysis. Under this framework, both learning and knowledge are key elements for changing the economic system (Malerba, 2005). The SSI framework is limited in focus to firms, assessing how established firms operate within a technology system over time. As such, SSI primarily covers incremental innovation processes (Malerba, 2005).

Malerba (2002) describes the innovation process as "a set of new and established products for specific uses, and a set of agents carrying out activities and market and non-market interactions for the creation, production, and sale of those products." Because the SSI approach to innovation is very product driven, it fails to capture many of the actors that operate alongside of established product firms, and has difficulty in considering the development of other networks or how core networks may transition over time.

As previously discussed, I have chosen a broader definition of innovation, in which the flow and creation of knowledge is important, as opposed to derivative, physical embodiments of this information. While the methods promoted through the SSI framework are well-suited for assessing incremental or accumulative innovations, like the efficiency changes that might take place *within* a sector (gradual electrification of vehicles), it fails to capture dynamic shifts *across* sectors to novel technological landscapes (automotive manufacturers like Tesla and the associated supply chains), as would occur through a creative destruction process (Coenen & Díaz López, 2010).

According to Malerba (2005), the evolutionary approach utilized within the SSI framework relies on three economic processes for driving innovation: (1) processes that create a variety of technologies, products, firms and organizations, (2) processes of replication, that generate inertia and continuity in the system, and (3) processes of selection, that reduce variety in the economic system (e.g. emergence of a few dominate electric vehicle manufactures). From this, the SSI framework is distilled to 3 main essential processes that need to exist to promote innovation:

1. Knowledge and technology (the actual technology)
2. Actors and networks (the people/organizations that employ and promote the technology)
3. Institutions (the policies and legal frameworks that promote or facilitate technology use)

Within this SSI framework, analysis beyond what happens within a sector, such as the societal patterns, actors, and the innovations that create a sector in the fist place, is limited. This can be illustrated by applying the SSI framework to the semiconductor sector as was done by Adams, Fontana, and Malerba (2013). Within this construction, it was recognized that different "user firms", firms that utilize the product or have demand for the product due to intrinsic values, acted as primary drivers for innovation within the sector. The case study presented below shows a process for innovation driven by an already existing sector and set of firms. These established firms continued to experiment with a specific technology, making incremental improvements to the process and supply chain over time. Missing from the narrative is the development of external networks and the information that accumulated to facilitate successful innovation outcomes, as well as the networks or mechanisms that existed to slowdown or deter innovation.

The case of semiconductor development within the SSI framework

The demand for semiconductors can be traced to the early markets that were established by the military, aerospace, and the computer industries (Langlois & Steinmueller, 1999; Malerba, 1985a). In the 1970s, new markets began to emerge in telecom, automobiles, and consumer electronics, which led to different and new applications for semiconductor technology (Malerba, 1985b). Throughout the 1980s and 1990s, two related factors directed semiconductor developments: the widespread use of semiconductors in mass consumer products, and the increased adoption of a new production process. These changes weakened the requirement that product design and manufacturing by handled by the same company, which ultimately facilitated the entrance of smaller firms, at both the design and manufacturing stages, into the market (Adams et al., 2013).

At the same time, firms such as Samsung and Philips, large firms that relied on existing chips to make a variety of different consumer electronics, gradually increased their demand for specialized chips for new applications. At the same time, it became increasingly more difficult to transfer the specific knowledge required to design customized semiconductor devices to the incumbent semiconductor suppliers. Rather than try to use chips that already existed in the market, or wait for semiconductor suppliers to design chips suitable for the required applications, these user firms gained access to the tools and knowledge necessary to design their own, customized chips in-house. Given the complexity in chip designs, user firms further had to interact directly with smaller, specialized semiconductor manufacturers, dramatically shifting the supply chain away from the integrated semiconductor suppliers (Glimstedt, Bratt, & Karlsson, 2010). Given this need for user firms to design custom semiconductor devices, alongside the reduction in production costs due to new production processes, user firms were able to eventually direct the path of semiconductor innovation (Adams et al., 2013; C. Brown & Linden, 2009; Ernst, 2005).

In this semiconductor analysis, Adams et al. assess the transition of the semiconductor industry over a period of time. Using the SSI framework, the transition is firm centric; it occurred because existing user firms demanded specialized products for specific applications. These specialized chips did not exist in the market, and so firms developed in-house knowledge and expertise to implement the necessary technology improvements and modifications. This case study portrays an instance where innovation was driven by already existing firms acting within an established sector. Firms like Samsung and Phillips took an existing product (semiconductors) and incrementally improved upon the product to meet their needs without developing revolutionary new technologies, or systems to bring about this change.

Through the SSI framework and the semiconductor case study, the demand for a novel product can be recognized as a primary driver of innovation. The SSI innovation narrative, however, only considers the demand side from specific user firms, and fails to explore the network of actors or external economies that came into existence to directly support or motivate these firm-level innovations in the first place. The SSI approach does not consider, for instance, why or how the demand for consumer electronics changed such that it required customized chip applications, nor does it consider the policies or institutional factors that may have affected this transition. Similarly, limited consideration is given to the incumbent industry, and the ongoing effects or adaptations that took place within that industry to accelerate, impede, or inhibit semi-conductor innovation over the same period of time.

Rather than assessing how two or more innovation networks shifted overtime, or came into existence, or dropped out of existence, the SSI framework instead focuses on the incremental incorporation of knowledge and information into an already strong and established network. While SSI is a useful framework for thinking about some innovations, this restricted view is best used for assessing incremental innovations within already developed networks, as opposed to assessing innovations that fall into the “creative destruction” category. Innovations that are disruptive, by definition, are innovations that manifest from new and different networks, ultimately displacing the previously dominant network.

Given the need to transition from a society that is heavily reliant on fossil fuels to a low-carbon one, a radically new network of actors and institutions will have to form. In turn, this network will need to supplant the existing fossil fuel network. This transition, to me, appears to be radically different from an incremental approach to innovation. Although some low-carbon innovations may ultimately follow incremental pathways (e.g. increased vehicle hybridization leading to fully electric vehicles), it is also likely that "creative destruction" will occur (e.g. fully autonomous vehicles that radically alter our current transportation system). Given this need, I think that alternative innovation frameworks may better capture the elements required to support a low-carbon technology transition.

1.3.2 Social Technical Systems

The ST-Systems framework consists of a multi-level perspective for innovation, in which innovation systems and regimes are established through a set of actors, institutions, and social frameworks, operating at different levels, to build a technological landscape (Figure 1.2). At the lowest level, technologies are experimented with and pulled into the market through the use of market niches, which act as test beds and engines for change (Geels, 2002). This framework relies on the establishment of initial niches and provides descriptions for the formation of technological landscapes as technologies move out of niche use into society at large.

ST-Systems explore the development of niches over time, recognizing that as individual niches build up, there will be spillover effects to different levels of the technological landscape. As these spillovers take place, new niche markets will be enabled due to links within the technological regime. This eventually may create a virtuous cycle that brings technology into the market in the form of a comprehensive technological landscape.

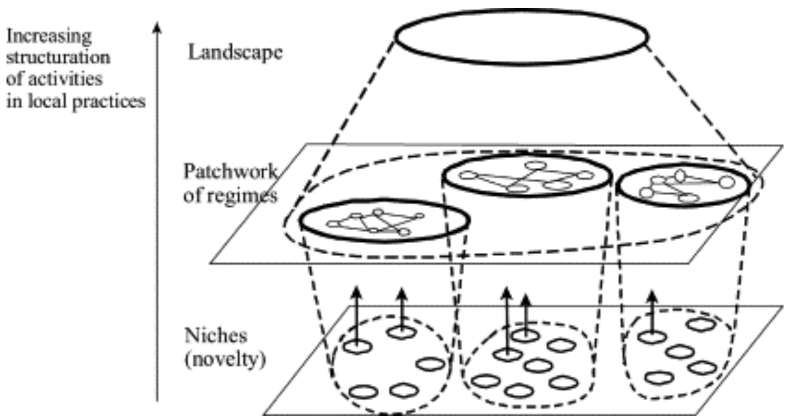

Figure 1.2. Multiple levels as a nested hierarchy (Geels, 2002)

The ST-Systems framework can be illustrated through the adoption of the steamship as portrayed by Geels (2002).

The case of the steamship within the ST-Systems framework

Geels (2002) characterizes steamship adoption in three distinct phases: the emergence of steamships in shipping, the period of sailing clippers and steamship use for passenger transport, and the phase of competition between steamships and sailing ships in freight.

The first phase began during the 18th century when shipping was surrounded by uncertainty and irregularity, with ships leaving ports when they were full and arrival depending on weather conditions. Shipbuilding was seen as an art, rather than a science, and involved a set of skilled laborers relying on intuition rather than calculations. The British, with a shipping monopoly, encouraged the production of a single ship-type thereby limiting design innovations. Innovations were only truly adopted after American Independence from Britain when American innovation led to faster ship designs (clipper ships).

As merchants and professional ship owners became increasingly prevalent, a secondary shipping market emerged reliant on contractual arrangements and the certainty of goods being carried between regions. The functioning of the shipping regime began to change due to institutional innovations associated with increasingly more dense networks of specialized middlemen (financiers, brokers, etc.). Innovations were made to promote the circulation of information (like journals of prices or commercial newspapers) rather than just goods, which led to a new niche market for shipping (Beniger, 1986).

The first intended steamboat application was a steam tug to pull ships through canals and harbors, first used in America in 1807, and Britain began a commercial steamboat passenger services shortly after in 1812. The steamer niche was limited to places where there was large-scale passenger and mail traffic, supplemented by special low-volume high-value cargo.

Eventually, the steam engine found a new niche when used as an auxiliary add-on to sailing ships to provide additional power when there was limited wind. From the mid-1830s onward the shipping regime changed as trade expanded, stimulated by relaxation of the Navigation Laws, and economic liberalization in Britain. British mail subsidies stimulated the use of steamships for mail transportation, and a global network of steam companies was created. The British mail subsidies created a protected environment allowing for the use and development of oceanic steamships, encouraging shipbuilders to emerge that specialized in the design and innovation of steamships. The use of oceanic steamers also added additional functionality in shipping by allowing for fixed departure and arrival times (Geels, 2002).

During this phase, experimentation began to increase the strength of boilers to provide more power and higher efficiency. This resulted in heavier steam engines, in turn leading to iron shipbuilding and additional spillover effects.

The second phase of the steam engine is characterized by developments from the 1840s through the 1870s, where a slew of social and economic conditions, like the Irish potato famine and the gold rush in California, led to shifts in technological landscapes and new steam engine applications. Transatlantic transport of passengers provided the first major market niche for steamships; rich emigrants were willing to pay extra for the advantages that steamships offered, and after the mid-1850s steamships quickly captured the majority of the emigrant market (Geels, 2002).

As iron ships became increasingly more prevalent and were adopted by the Navy, incremental improvements in boiler technology took place to promote higher efficiency and increased speed. The introduction of compound engines made it possible to use steamships in particular long-distance market niches in freight shipping. In 1866, steamers with compound engines, using 40% less fuel, competed successfully in the China tea trade with sailing ships (Geels, 2002).

The third phase for steamship adoption is defined by increased competition with the incumbent sailing ship market. Emigration to America sped up through the 1880s and

1890s, providing a profitable market niche for steamships while expansion also occurred in freight transportation. The opening of the Suez Canal in 1869 not only shortened distances to the east, but was unsuitable for sailing ships due to few and variable winds.

Nevertheless, the diffusion of steamships was gradual and sailing ships continued to be used well into the 20th century. Steamship adoption was hampered because knowledge and technology centered on improving engine efficiency and material integrity took time to develop. Additionally, sailing ships turned to efficiency improvements of their own by increasing ship size and lowering the cost of freight transport. Metalworking also improved which further lowered the cost of shipbuilding, and improvements in rigging decreased labor requirements for sailing ships. Eventually, incremental improvements in the steamship design led to lower fuel costs, and decreased costs for freight transportation. As sailing ship markets continued to erode and port and harbor sizes increased to accommodate ever-larger steamships, the steamship became ubiquitous.

From the above narrative, it is clear how the ST-Systems framework is applied. A technology is spread through one market niche after another, where different advantages of that technology are realized, before it is able to compete with, and eventually displace a previously dominant technology. While the ST-Systems literature provides a robust framework for describing the adoption of technology through ever-prolific niche markets and provides some narrative on the creation of agent networks, it does not offer a dynamic view of the agent network (how it shifts). The ST-Systems framework also does not fully consider how the blocking and inducement mechanisms may deter or promote the formation of specific networks (Coenen & Díaz López, 2010). To better understand why technology uptake can occur or not occur in specific niche markets, and to understand how to foster the development of new niches, a different approach is still necessary. Of the frameworks I have looked at, I find that the Technology Innovation Systems framework is better suited for exploring these dynamic processes.

1.3.3 Technology Innovation Systems

The TIS framework is similar to the SSI framework: it is founded on principles of evolutionary economics as a means of describing technological innovations (Bergek & Jacobsson, 2003; Carlsson & Stankiewicz, 1991). The difference from the SSI framework, however, is that TIS focuses on the network of actors and the dynamics of developing an innovation system that promotes innovation, as opposed to assessing individual innovations that occur strictly at the firm level.

Within the presence of an entrepreneur, and when a given technology network has received sufficient critical mass, the network may be transformed into a development block – or a synergistic cluster of firms and technologies operating within an industry or group of industries (Carlsson & Stankiewicz, 1991). This transformation is similar to what is described by the ST-Systems framework in relation to niche market proliferation. The TIS framework deals primarily with flows of knowledge and competence, and the networks that develop to use this knowledge over time, rather than just the flows of ordinary goods and services (Carlsson & Stankiewicz, 1991). Because the TIS framework considers network dynamics, it offers opportunities to better explore technology transitions, and how different networks come into existence, or falter over time. TIS provides a frame for assessing how policy intervention may help direct or assist technology transitions.

The work of Bergek, Jacobsson, Carlsson, Lindmark, and Rickne (2008) has further refined the innovation system put forward by Carlsson and Stankiewicz (1991) to develop a set of seven functional forms that exist within an innovation system (Table 1.2). These functions were developed through use of previous literature on innovation and through expert opinion on the innovation process (Bergek et al., 2008).

Table 1.2

The seven innovation functions (Bergek et al., 2008)

1. Knowledge development and diffusion
2. Entrepreneurial experimentation
3. Influence on the direction of search
4. Market formation
5. Development of positive external economies
6. Legitimation
7. Resource mobilization

These 7 functions are useful for understanding what specifically occurs in the process of innovation. Each function represents something that happens in the innovation process. Actions can be taken that either support a function, or that block a function. The relative level of support or hindrance ultimately affects the innovation network, and path of innovation. If any given function does not receive sufficient support, innovation is ultimately deterred.

In turn, it is possible to use these functions as a means to map existing actors on to an emerging technology. This provides insight into the development and the trajectory of the innovation system. For instance, if a given function has received minimal support within the TIS, or there are a number of blocking mechanisms (e.g. lawsuits are preventing an advanced biofuel mandate from being enacted, detracting from market formation), innovation and development is likely to be deterred. Assessment and discussion of innovation system weakness has been demonstrated in the case of the Swedish wind turbine industry (Bergek & Jacobsson, 2003). In Sweden, diffusion and adoption of wind turbines was hindered due to inadequate functional fulfillment; there was inadequate knowledge development for wind turbines while market formation support was strong due to implemented policy. This weakness in functional fulfillment ultimately lead to the deployment of lack-luster technology lacking in legitimacy.

Unlike the SSI framework that focuses on firms, or the ST-Systems framework that focuses on the use niche markets, the TIS framework focuses on the development of actor networks, which may include firms and niche markets. This distinction can be seen in the case study of the Dutch biofuel innovation system, in which technology development and innovation is assessed through a series of events that link back to a set of functions analogous to those found in Table 1.2.

The case of Dutch biofuels within the TIS framework

Suurs and Hekkert (2009) note that there are two distinct technology classes within the Dutch Biofuel Technology Innovation System (BFTIS). These technology classes consist of first generation biofuels (1G) and second generation biofuels (2G), where the 1G fuels originate out of the agricultural TIS, and the 2G fuels stem from bioscience-based technologies.

The biofuel technology case traces its origins to agriculture, where massive crop production led to a surplus of food supply in Europe and decreased revenues. As a means to support the agriculture industry, Europe instituted a number of generic tax exemptions that influenced the direction of search toward biofuels. Farmers were offered a premium for the cultivation of non-food crops to promote biofuels as a new market for agriculture products.

The first phase of the Dutch BFTIS began to take shape when a group of entrepreneurs in the Netherlands started using biofuels, making use of their own funding as well as European subsidies ("resource mobilization"). Despite this initial market, biofuels had low economic feasibility and were unable to compete with fossil fuels. At this time, the Dutch government provided no support for biofuels, and the government agency for energy (Novem) was against the use of biofuel, viewing it as too expensive compared to co-firing biomass in power plants (Suurs & Hekkert, 2009). This hindered the legitimacy, and influenced the direction of search away from biofuels. Although Novem expressed doubt about biofuels, the Dutch Ministry of Agriculture favored biofuel development; these contradictory positions from different government departments led to market uncertainty (detracting from "direction of search"), which was compounded by a lack of monetary support ("resource mobilization").

The second phase of the Dutch BFTIS was shaped through entrepreneurial experimentation, knowledge development and diffusion, and through public opinion and industrial actors influencing the direction of search toward biofuels. Climate issues also became a matter of political interest ("exogenous external economies"), and biomass was recognized as an important consideration in the energy sector.

Two boating companies experimented with biodiesel use at this time, later leading the companies to demand a national fuel tax exemption for the project, which the provincial government and the district board of agriculture supported ("influence the direction of search", "resource mobilization", and "legitimacy"). A tax exemption was provided, and a virtuous cycle emerged when several other boating projects started that also demanded tax exemptions (Suurs & Hekkert, 2009).

In 1995, Nedalco, an alcohol producer, along with other connected companies pressured the national government to change the tax scheme and issue a tax exemption for an ethanol production facility; a tax exemption for the annual production of 30 million liters of bioethanol was provided in 1997. Although the tax exemption ended up being insufficient to cover facility investments and the project was discontinued, Nedalco successfully eroded the Dutch government's resistance to (1G) biofuels (Suurs & Hekkert, 2009).

The third phase in the Dutch BFTIS is marked by the creation of a carbon-neutral energy carrier program (GAVE) initiated by the Dutch agency for energy and motivated due to climate change concerns. An influential study authorized by GAVE indicated that biofuel production could be favorable and that a range of alternative energy sources already existed for electricity production ("legitimacy", "influence the direction of search"). This argument led to a national agenda for alternative fuel development which strongly favored 2G technologies, resulting in the development of two major industrial coalitions – one spurred by Nedelco and another by TNO, Shell, and Wageningen University (Suurs & Hekkert, 2009). The promise of 2G technologies created positive BFTIS dynamics at the same time the negative aspects of 1G biofuels were being stressed by academics and environmentalists, stagnating entrepreneurial experimentation and knowledge development of 1G fuels.

The fourth phase in the Dutch BFTIS was marked by additional support from the GAVE program aimed at guiding entrepreneurs toward demonstration projects. Due to a limited budget, however, commercial facility plans were discontinued and the subsidy programs stopped; the absence of sufficiently powerful market creation policies formed a critical barrier for further development of the BFTIS (Suurs & Hekkert, 2009).

The fifth phase was marked by European intervention triggered by the 2003 EU biofuel directive, a directive that forced EU member states to substitute a percentage of automotive fuels for biofuels ("market formation"). As the EU initiative did not dismiss 1G biofuels, a new wave of 1G promotion in the Netherlands occurred. Nedalco continued to influence the field and drafted a proposal for a large-scale 1G bioethanol facility. Despite the lack of a national policy to promote biofuels, several other 1G fuel projects supported by various corporate coalitions commenced, triggered solely by the EU initiative. The controversy around 1G and 2G continued to increase, yet the choice for 1G or 2G biofuels, which was first presented as a dichotomy, became irrelevant in the BFTIS.

The final phase studied by Suurs and Hekkert (2009) is characterized by a slew of Dutch policies mandating the use of biofuels and providing R&D support and subsidy for 2G biofuels. At the same time, the resistance to biofuels (both 1G and 2G) increased, and biofuel use remained surrounded by controversy associated with land-use for energy crops, rising food prices, and the deforestation of vulnerable natural areas like rainforests. Additionally, potential 2G biofuel producers were deterred from producing due to the uncertainty in the biofuel market as it remained to be seen whether or not 2G biofuels could compete with 1G biofuels, which already had issues competing with cheap biofuel imports and conventional fuels.

The event sequence related to 1G biofuels is characterized by market formation, further encouragement in the direction of search, and additional entrepreneurial experimentation and resource mobilization to support those efforts. The use of 2G biofuels in the Netherlands was only driven by entrepreneurial experimentation, plagued by considerable uncertainty due to the rapid expansion of 1G biofuels and controversy over the use of biofuels in general, which undermined the long-term perspective for development of biofuel technologies.

This actor-network narrative for the Dutch biofuel case provides a compelling story for the shift of a technology sector over time, and allows for incorporation of firm-based innovations and niche market strategies as part of the discussion. Given this unique characteristic, the TIS framework is effective for exploring innovation system transitions and for assessing the role that policies may play in fostering transformational technological shifts.

All of the frameworks for technology adoption that have been discussed above have a set of similarities: namely that innovation is an ongoing process that requires a network of different actors and agents diffusing and utilizing information to progress. With this framing in mind, I turn to the body of work for forecasting innovation.

1.4 Forecasting Technology Innovation

Economists have traditionally looked at technology innovation from a macroeconomic view to better understand how innovation drives economic growth. Engineers, on the other hand, are interested in the application of innovation to specific technologies, and have the need to forecast improvements that can be expected to occur for technologies over time. I now turn to the engineering side of the problem – bottom-up approaches that explore technology innovation.

Technologies in their earliest versions are often costly and provide limited use. To achieve widespread deployment and market uptake, these technologies require large reductions in cost alongside performance improvement. The previously presented steamship case study perfectly exhibits this process. Many important technologies have experienced a process of steady and consistent cost reductions and performance improvements over time. Semiconductors, cellular phones, photovoltaics, and gas turbines have all experienced innovations that have decreased costs, facilitating increased market penetration and use. This relationship between cost and deployment is captured by "experience curves," a mathematical formulation that illustrates a simple, quantitative relationship between the price of a technology and its cumulative production or use (IEA, 2000).

Wright (1936) first introduced the concept of technological learning, where he established that the labor hours required for constructing an airframe decreased with increasing production. In 1962, the concept of learning-by-doing was introduced to endogenous growth models (Arrow, 1962). A variety of studies have since found support for the use of experience curves to forecast innovation (e.g. Alberth, 2008; Rubin, Yeh, Antes, Berkenpas, & Davison, 2007; Yeh & Rubin, 2012).

Forecasts for technology costs and adoption frequently rely on this well-established correlation between scale and cost, and this simplistic relationship often factor into the engineering-systems literature to forecast price improvements for specific technologies (e.g. IEA, 2000; Söderholm & Klaassen, 2007). For low-carbon energy technologies, there has been increased reliance on the experience-curve model. Today's deployment policies and integrated assessment models (IAM) rest on the assumption that a steady decline in technology costs can be achieved mainly by boosting installed capacity or unit production (Loulou, Kanudia, Lehtila, & Goldstein, 2005; U.S. Energy Information Administration, 2012; van der Zwaan, Gerlagh, G, Klaassen, & Schrattenholzer, 2002). As a result, these policies and models often focus on an approach dubbed "riding down the experience curve" for reducing the costs and improving the competitive advantage of low-carbon technologies compared to fossil incumbents (IEA, 2000; van der Zwaan et al., 2002).

In the best-case scenarios, IAMs incorporate "endogenous technological learning" that makes use of experience curves to shift the equilibrium state of technological deployment. These simplistic learning curve models assume costs will decrease a certain percent for each unit of additional production. There are two important weaknesses with this approach: (1) the rate of the decrease is highly uncertain, and may not follow historical rates for similar technologies, and (2) these models assume that technologies are already at a commercial stage and are capable of following a standard learning trajectory (Lohwasser

& Madlener, 2013; W. Nordhaus, 2009). In reality, however, the progress ratio (rate at which technology costs decrease per unit of deployment) is variable, and experience curves become increasingly less accurate as the magnitude of deployment increases (Alberth, 2008; van Sark, Alsema, Junginger, de Moor, & Schaeffer, 2008).

Given the weaknesses of these "single-factor" experience curves, the model has been expanded, in some instances, to include "two-factor" learning (Lohwasser & Madlener, 2013). Two-factor models add a parameter for R&D, which is used to overcome the bias for early-stage technologies (Yeh & Rubin, 2012). These "two-factor" models try to incorporate knowledge development, or a proxy for how much knowledge about a technology currently exists. To add knowledge stocks to the model, patent counts are often used as a proxy to establish how mature the technology is (Jamasb, 2007; Kouvaritakis, Soria, & Isoard, 2000).

The experience curve model of the innovation process is specifically concerned with cost reductions, assuming that a decrease in cost is the desirable outcome of innovation, ultimately leading to a positive feedback loop which generates expanded deployment and further reductions in cost. Various deployment policies are based on this model, assuming that mandated deployment would eventually reduce costs for low-carbon technologies. Renewable electricity mandates, tax incentives and feed-in tariff policies, low-carbon and renewable fuel standards, and even cap-and-trade and carbon taxes are all assumed to drive down the costs of clean energy sources over time by spurring deployment and accelerating experience curves.

However, these policies are overly reliant on the assumption that deployment alone is enough to drive reductions in cost, and do little to ensure that costs follow the trends associated with the experience curve (Jamasb, 2007). While the correlation outlined in experience curves tends to hold true across a variety of industries, policies designed to bluntly accelerate experience curves ignore the underlying and interrelated causal factors of technological change at work. For instance, policies have been effective at accelerating the deployment of wind turbines, but costs have not declined as predicted by historic progress ratios, and have gone up in recent years (IPCC, 2011). Similarly Germany's Feed-in Tariff policy has resulted in substantial deployment of rooftop solar, but has not resulted in reductions in cost that follow historic progress ratios for photovoltaic technology (Wand & Leuthold, 2011). Missing from the experience-curve approach is the nuance of actor-agent networks and a systems view of innovation. Instead, the use of experience curves rests on the assumption that deployment is a causal driver of price reduction, and that government-supported deployment in and of itself is sufficient for driving innovation.

As a result of the over-simplification to these innovation models, past and current energy policies over-rely on scale to achieve cost reductions while routinely ignoring the near-ubiquitous role of the other factors necessary to optimally incorporate learning feedbacks and the benefits of experience. Technology deployment policies reliant on experience curves therefore overestimate the technological change resulting from deployment alone and risk continually underperforming cost reduction objectives (Schmalensee, 2015).

A different modeling approach for the adoption of innovative technologies and innovation takes shape in the form of the diffusion curve (Meade & Islam, 2006). Diffusion curves have traditionally been modeled as S-shaped-curves, following a standard logistics regression (Figure 1.3). One possible explanation for this diffusion pattern is that different markets (or individuals) have need for the technology, or are curious about the technology in its nascent stages (Meade & Islam, 2006). The first adopters of a given technology are often called the “innovators.” Innovators are defined by their willingness to take risks in technology adoption, deriving some benefit (either intellectual curiosity, or otherwise) from a technology that is likely inferior and more expensive than a similar technology that already exists on the market. There has been some evidence to support the idea that innovators are wealthier and that diffusion to the masses is driven by income and technology price (Karshenas & Stoneman, 1992; Wareham, Levy, & Shi, 2004). As the technology matures, early adopters begin to purchase and experiment with the technology. As the technology is further utilized, diffusion starts to increase until it is commonplace, and a large subset of the maximum theoretical market is employing the technology. Even at this stage, there is still some resistance to adoption – this market segment that resists adoption is termed the “laggards.” Figure 1.3 shows the logistic curve used to represent typical diffusion.

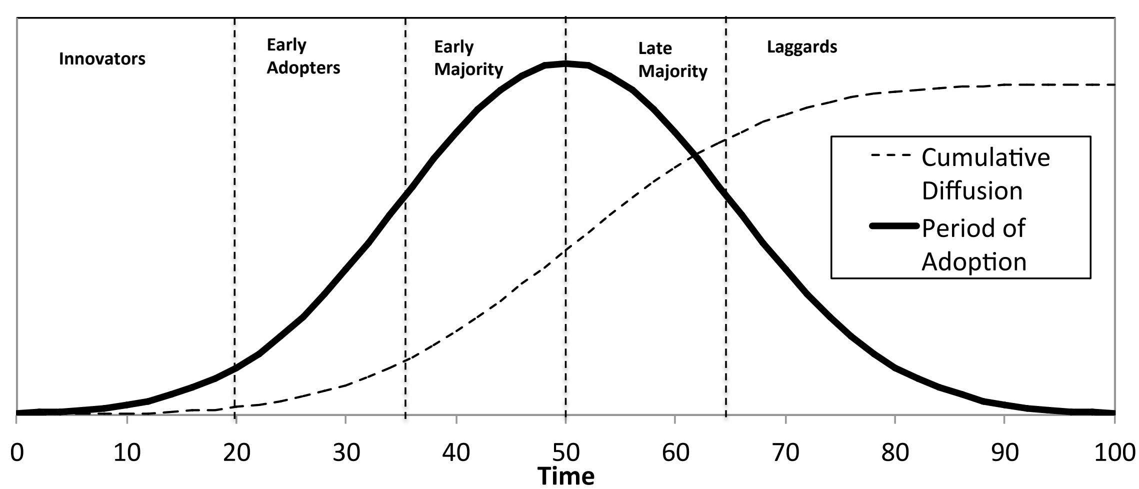

Figure 1.3. The standard model for diffusion. Innovators or 2 standard deviations or more above the mean.

Models that make use of these curves often provide exogenous adoption timeframes, and use the curve to inhibit product uptake independent of price (U.S. Energy Information Administration, 2012). While limiting uptake in this fashion is not entirely baseless, the time-frame for adoption depends widely on technology specifics, as well as other socio-technical factors (Van den Bulte, 2000). Numerous modifications and improvements have been made to the standard diffusion model to try to better account for adoption rates. These modifications and developments range from considering diffusion in two or more stages (Kalish, 1985), to examining the deterrents to adoption and the role of advertisement (Horsky & Simon, 1983). Increasingly more complex model formulations and parameters have been added to the diffusion model to try to improve model prediction of ultimate technology adoption (Meade & Islam, 2006).

Despite improvements in diffusion modeling, there still remains uncertainty associated with the parameters and deterrents of diffusion especially when limited data exist. Furthermore, when diffusion of technology ultimately depends on displacing an incumbent technology, and allocating scarce resources to achieve this result, the diffusion models can offer only limited guidance.

In terms of evaluating the economic tradeoffs between traditional R&D expenditures, additional advertisement, improvements in technology costs through economies of scale, or further demonstration and legitimation of the technology, today's existing models remain lacking. Missing from these models is the need for a more nuanced, system view of the innovation process.

1.5 The Need for Better Approaches to Innovation

Both the top-down macroeconomic models and the bottom-up engineering approaches to innovation have been used to motivate innovation policy. These innovation policies seek to influence energy technologies in two, disconnected ways that may ultimately undermine policy efficacy and societal goals. One of the methods employed to promote innovation is through the use of deployment policies to accelerate adoption. Deployment policies often target and favor more mature technologies but are poorly optimized to drive down costs or promote non-incremental innovation. The second method is through public funding for R&D efforts. These efforts focus on nurturing nascent technologies with no consideration of their commercial applications. The resulting R&D and deployment policies are fragmented, inadequately supporting and accelerating many of the underlying technological, industrial, and economic factors necessary to drive the feedback loops that bring down technology costs and drive innovation (J. Jenkins et al., 2012; Sagar & van der Zwaan, 2006).

Similarly, a new approach is necessary to break into the "innovation black box" for both the engineering and macroeconomic innovation models. These models could greatly be improved by better capturing the effects that drive and promote technology innovation. A scalable approach is necessary, one that can be employed across regions, considering a myriad of technologies with different costs, while building on existing innovation theory to better assess, work with, and understand innovation. From this need, my work hopes to provide tools that can be readily applied to help break into the innovation black box. It is my goal that the methods and data employed here may help to discern what is happening with a technology, and can be used to aid policy makers in determining where scarce resources may be better allocated to promote specific technology innovation goals going forward.

Chapter 2: Methodologies for Building on the Innovation Systems Framework

As innovation studies continue, there is hope to eventually break into the innovation black box, elucidating the levers and mechanisms that can be turned, twisted, and caressed to yield desirable economic outcomes. Despite the recognition that innovation and the innovation process is important to both assessing technology uptake, as well as economic systems at a macro level, approaches capable of better quantifying the innovation process have been limited.

Bottom-up models of innovation make use of diffusion curves and experience curves, simplistic models with high uncertainty that fail to account for the drivers of innovation beyond cumulative deployment and R&D investment. Top-down models similarly struggle with innovation, failing to endogenize all aspects of the innovation process. It is therefore unsurprising that many innovation system frameworks have emerged, facilitating qualitative assessment of the innovation process at the micro-level, providing a means for better conceptualizing innovation.

It is these innovation frameworks that I find to be the most promising. My research therefore builds on the Technology Innovation System framework for innovation. I aim to create a set of tools and methods that can be employed to better inform policy makers, modelers, and stakeholders about the innovation process, directing them toward desirable policy approaches that support the transition to a low-carbon society.

For my research, I use a broad definition of innovation, which I initially established in chapter one: the creation, dissemination, and use of new information. My approach to understanding and predicting innovation outcomes draws on the idea that innovation can be tracked and understood by analyzing the flow of information relating to specific technologies. Schumpeterian entrepreneurs rely on information flows to create, diffuse, and utilize novel technologies and groundbreaking innovations. In turn, these very actors contribute to these information flows, strengthening the network, and further supporting innovation. This conceptualization emerges from the TIS literature, which focuses on the idea that the structural arrangement (specific actors) of the innovation system emerge when adequate support is provided for the levers (TIS functions) that serve to promote these structural components (Bergek et al., 2008).

Despite the quintessential role that information plays in driving entrepreneurship and fostering innovation, modeling approaches usually fail to account for the importance of information. Even the best models apply only a superficial overlay of "knowledge stock proxies" (e.g. Popp (2004)).

As discussed in chapter one, innovation relies on networks of agents that have access to information and that are capable of utilizing the information. I posit that the integration of innovation into models and policy discussion can be improved by tracking the flows of information that are available across different networks. It is important to consider not only the quantity of information being produced, but also how the information is accessed, how readily available it is, and the quality and type of that information.

It is impossible to obtain and assess all the information associated with a given technology, and it is even harder to establish the set of actors that may use or build on that information, and how they come into contact with it. Nonetheless, there are certain informational sources that may serve to capture part of the available knowledge base.

Patents represent but one type of data believed to capture part of the information flow that drives innovation (Agrawal & Henderson, 2002; S. Breschi, 2001). In addition to the use of patents, I turn to newspaper articles and trade journals, as well as in-person meetings to better assess different information flows. While incomplete, these informational resources provide some insight into the types and quantity of information relevant to technology. As such, these resources can act as possible proxies for technology innovation.

Due to the heavy reliance on patents as a proxy for innovation in the literature, I also utilize patenting activity to assess innovation. I further use this conventional innovation proxy as a baseline for comparing other possible innovation proxies. To establish this baseline, I look at patent activity for one of the most mature alternative fuel technologies: biofuel.

After using patents to establish the baseline of innovative activity for biofuels, I explore the use of newspaper data as an alternative proxy for technology innovation in Chapter 5. In turn, I make direct comparisons between the article innovation proxy and the patent data associated with biofuels. Building on these results, I assess the use of the article innovation proxy for an emerging technology field: electric vehicles. I further corroborate these results through qualitative research methods, namely in-person meetings in several states in the U.S., to develop a more complete picture of the technology innovation narrative for electric vehicles.

To better characterize and understand the information I have gathered, I draw on computational approaches to information theory. I use these approaches for two main purposes: (1) to better classify and group patents together to assess how patenting activity for specific types of technologies has changed over time, and (2) to better classify and sort the information I gather into innovation functions, or the specific kinds of events and actions that occur to guide the structural development of the innovation system in specific ways. To achieve these classifications, I use natural language processing (NLP) and apply it to large bodies of text.

NLP is a technique that uses computational approaches and statistical models to analyze text for the purpose of achieving human-like language processing (Liddy, 2001). NLP is based on a set of theories and technologies, and remains an active, ongoing area of research (Cambria & White, 2014). NLP offers the advantage of being able to search through large bodies of text for users seeking to paraphrase, translate, or answer questions about the content of the text. While NLP systems are not able to directly draw inferences from the text, NLP can provide the user with the necessary tools and results to better or, more simply, draw meaningful inferences. For my research, I extensively use the Stanford NLP Classifier (available from: http://www.nlp.stanford.edu/software/classifier.shtml).

The NLP approach that I rely on is called "supervised learning." This approach is used to train machine-learning algorithms for the purpose of classifying texts. Classifying text through a supervised learning method is accomplished through a two-step process: the first step is to develop the classification model by creating a "training" set of texts, and the second step is to classify other texts using the "trained" model (Grimmer & Stewart, 2013). To train the model, a small subset of manually classified texts is utilized. These texts are read in full, and then sorted into one or more user-defined classifications. From these user-defined territories, the NLP algorithm is able to break each text down into a set of components and features, which can then be used to create a statistical model to predict how new texts should be classified.

Additional parameters are provided to specify how the text should be parsed and broken down into its component elements. This may include things such as article length, word pairs (e.g. “the hat” or “electric vehicles”), and stemming properties (“run” versus “running”). The model is then optimized to improve accuracy and prediction for text classification within the training set.

After the text-classification model has been trained, the remaining set of texts can be tested against the model to provide insight into how accurate the model is at classifying new text documents.

2.1 Using Patents to Assess Biofuel Innovation

To establish an initial narrative for biofuel innovation in the United States context, a literature review was conducted prior to the collection of patent data. The Google Scholar electronic database was utilized to find relevant journal articles and government reports to provide historical context for biofuel developments. The following keyword search was initially used: “Biofuel + History + United States.” The electronic database returned thousands of results based on this search query.

To further refine the selection of relevant articles, results were sorted by relevance, and all article titles within the first 100 results were read in full to assess whether or not the article was relevant to the greater historical context of U.S. biofuel developments. All papers deemed to be relevant were downloaded and read in full. Additional historical context for the narrative was established by examining the works cited in the papers from this initial survey. Well-cited articles not found in the initial survey of literature, but that appeared in the citations of these initial publications, were also downloaded and read to aid in the construction of the biofuel narrative. Articles associated with biofuel development in the broader U.S. context were chosen as the key articles with which to construct the biofuel narrative in Chapter 3. Articles that were too narrow in scope, focusing only on a single biofuel technology without broader historical context, were omitted.

From the biofuel literature review, a biofuel narrative emerged. The validity of this narrative was assessed at several academic conferences through in-person meetings, as well as public presentations. Experts in the field offered many helpful suggestions. Using this approach, a "default" biofuel narrative has been established, which I have used to shape and facilitate the discussion and analysis associated with the patent data I collected.

Patent databases represent an appealing data source for tracking and assessing innovation. Patent data is publicly available, and patent systems have been instituted as a means to protect valuable intellectual property (Mogee, 1991). Patents are therefore a public record of invention – inventions that may ultimately be used and deployed. As such, patents provide one measure of the flow of information associated with technology innovation.

Patent analysis has been used to track trends in knowledge development and diffusion related to new technologies and inventions (Hekkert, Suurs, Negro, Kuhlmann, & Smits, 2007; Popp, 2005). Patenting behaviors can also reveal overall trends in entry and exit of firms in the market (if firms continue to file for patents or stop patenting), and ultimately the direction and maturity of innovation processes. Patent analysis, however, is not straightforward. It is difficult to assess the value or quality of any individual patent, and there are differences in patenting activity between sectors; this makes the absolute level of technology innovation ambiguous (Archibugi, 1992; Peeters & Pottelsberghe de la Potterie, 2006; Qian, 2007). Innovation also depends on a number of conditions that are not fully captured in the patent dataset. Patent data therefore capture an incomplete picture of the innovation process. Additionally, patenting does not necessarily indicate that anything useful has been created, and an increase in patenting could possibly retard innovation, as opposed to accelerate it (Heller, 1998; Moser, 2013). Furthermore, there is a long lag-time between the point that a patent is filed, and the point at which it is granted (Popp, Juhl, & Johnson, 2003). This lag-time also makes it difficult to use patents as a real-time measure of innovation.

Despite patent data limitations, I aim to assess the overall trends associated with patent filing for biofuels under the assumption that these trends may adequately represent significant shifts within the biofuel innovation system. Patents contain specific textual information concerning individual technologies. Using patent classification, it possible to obtain some disaggregation of the technology innovation process compared to aggregate data such as R&D expenditures (Popp, 2005). I look at the existing biofuel innovation narrative in the U.S. to establish if patent activity adequately captures many of the innovation system changes that have occurred. For instance, limited patenting activity followed by a sudden surge in patenting activity, regardless of the actual quality or value of any individual patent, may be indicative of substantial shifts in that technology's innovation system. Additionally, there are few publicly accessible datasets that can replace the use of patents as a proxy for innovation. However, when utilizing patents to assess knowledge stock, there is significant subjectivity in selecting classifications that relate to technologies (S. Breschi, 2001). For emerging or innovative technologies, this can be especially difficult if classifications cover the invention inadequately, or if the invention builds on a large body of research in another, related field. Rather than relying solely on the classification scheme established by patent examiners as the basis for technology classification, I employ NLP techniques to create a unique, technology-specific statistical model to aid in patent grouping and classification. I classify patents into 1 of 10 possible technology categories: 1^{st} and 2^{nd} generation ethanol and biodiesel, renewable diesel or drop-in fuels, other renewable fuels, Fischer-Tropsch synthesis not based on renewable feedstock, oil recovery, other non-renewables, and non-applicable (NA) patents. This approach can increase the precision and accuracy in tracking patent activity associated with individual technologies.

While assessment of patent quality is not a direct requirement for characterizing shifts in a technology's innovation system, further investigation into innovation system strengths or levels of inventiveness requires more consideration of the quality of patents.

I rely on three patent classification systems to assess patents as a proxy for shifts in the biofuel innovation system. I utilize (1) The Green Inventory for biofuel, a system based on a selection of International Patent Classification (IPC) codes, (2) the existing Cooperative Patent Classification (CPC) scheme for biofuel, and (3) my own derived methodology that makes use of natural language processing and machine-learning algorithms to independently classify patents. I used a multi-step process to collect a set of relevant patents from the United States Patent and Trademark Office (USPTO) for this analysis.

I collected patent counts for patents applied for in each year for CPC and IPC classification approaches. These counts were obtained from the USPTO using built-in search functionality to search by classification code. Both the CPC biofuel classification scheme and the IPC Green Inventory were created to capture a substantial portion of patents likely to be associated with biofuel technologies across a large portion of the biofuel supply chain. Table 2.1 provides a list of the patent classification identifiers used to construct the Green Index and CPC biofuel patent trends used for analysis.

Table 2.1

Patent Classification identifiers for BioPat patents and CPC biofuel patents

Technology	Green Index (IPC)	CPC Classifications
Biodiesel	C07C 67/00, C07C 69/00, C10G, C10L 1/02, C10L 1/19, C11C 3/10, C12P 7/64	Y02E50/13
Bioethanol	C10L 1/02, C10L 1/182, C12N 9/24, C12P 7/06-7/14	Y02E50/17
2G Alcohol	NA	Y02E50/18
Cellulosic Ethanol	NA	Y02E50/16
Pyrolysis	NA	Y02E50/14

Note: C07C relates to organic chemistry and cyclic or carbocyclic compounds; C10G relates to the production of liquid hydrocarbon mixtures and cracking hydrocarbon oils; C10L relates to fuels not otherwise captured in C10G and C10K; C11C covers fatty acids obtained from fats, oils, or waxes; C12P relates to fermentation, and C12N relates to microorganisms. Conversely, Y02E50 covers to the production of fuels of non-fossil origin with /13 covering biodiesel, /17 covering grain bio-ethanol, /18 covering bio-alcohols produced through methods other than fermentation, /16 covering cellulosic bio-ethanol, and /14 covering bio-pyrolysis.

To establish an initial repository of patents likely to be associated with biofuel innovation for my own classification methodology, 126 biofuel-relevant keywords were utilized to search the USPTO patent database. This keyword list was originally developed as part of the Biopat database research conducted by Costantini, Crespi, and Curci (2013) and is contained in the appendix. These keywords were established by Costantini et al. (2013) using expert elicitation and a review of relevant biofuel research literature. This keyword set is believed to describe adequately capture the technologies and processes employed across biofuel production. These keywords are often very broad, including words such as “ethanol,” an industrial solvent and chemical with a myriad of uses and applications outside of biofuels.

The combined keyword search of the USPTO database returned more than 2.4 million patents, capturing a greater subset of possible biofuel patents than covered under the BioPat database (47,500 patents) or the CPC classification scheme (3,300 patents). This approach generated a repository similar in breadth to the World Intellectual Property Organization’s (WIPO) IPC Green Inventory for biofuels. Accounting for duplicate patents[1] in the patent dataset yields a full patent dataset to assess for biofuel-relevance of over 755,000 patents.

To better assess patent trends and meanings, it is important to disaggregate and properly classify relevant patents (Popp, 2005). Because keyword searches capture a number of patents that have no relation to biofuels, a more robust method for classifying biofuel patents is necessary. I have chosen to utilize natural language processing to

[1] Patents may have more than one keyword associated with them.

augment the other biofuel patent classification schemes (Adelman & Deanglis, 2007; Allison & Lemley, 2000).

The patent text and associated assignee information and patent classifications for each patent was downloaded in full and retained in a local patent database. To create a classification model for biofuel patents, 1000 patents were randomly selected from the patent database. Each of these patents was read in full, and then was manually sorted into 1 of 10 categories based on the technology use implications.

Patents classified as relating to ethanol or biodiesel were also classified as being either 1st or 2nd generation technologies based on whether or not the technology made direct use of conventional food crops. Technologies making use of non-food or non-conventional resources were sorted into 2nd generation technologies, which included technologies like cellulosic ethanol production from corn stover, biodiesel production from algae, and thermal chemical conversion processes.

For manual classification, I did not differentiate between process-oriented patents and feedstock-oriented patents, and instead classified patents based on their relation to the overall biofuel production processes. Process technologies for fuels meant to be direct replacements to conventional gasoline and diesel were sorted into the renewable diesel or drop-in fuel class, which I include in the 2nd generation technology classification. Patents relevant to thermochemical production of biofuel were only classified as biofuel patents when biomass-related feedstock or applications were directly mentioned or cited within the patent.

Utilizing the Stanford NLP Classifier, I created a model for classifying all 755,000 patents in the patent database. The model was built using 700 randomly selected patents from the initial repository of 1000 manually classified patents in the patent database (Kessler, 2015c). The model was then assessed for accuracy using the remaining 300 manually classified patents that were not used to create the model. Model accuracy for each category varies, with overall classification improving for technologies with a greater representation in the random sample.

Due to the small random sample of the total patents used for manual classification, not all categories have sufficient information for producing meaningful inferences. I have therefore aggregated model classification results into five main categories: not-applicable patents, 1st generation patents, 2nd generation patents, ethanol patents, and biodiesel patents. Table 2.2 shows the bin sizes and relative accuracy for the biofuel-relevant classifications.

Table 2.2

Validation results for biofuel technology classifications

Biofuel Patent Type	Correctly Classified	Not Classified as Biofuel	Total in test set	Total in training set	Classification Accuracy
1st Generation	121	10	131	148	92%
2nd Generation	27	24	51	42	53%
Ethanol	87	26	113	119	77%
Biodiesel	57	4	61	60	93%

The provided classification accuracy is based off of the number of patents correctly identified by the model as belonging to a specific class of biofuel patent in the test set compared to the total number of patents that were manually classified as that type of biofuel patent in the same test set. Figure 2.1 is visual representation of the different levels of data considered for this classification approach.

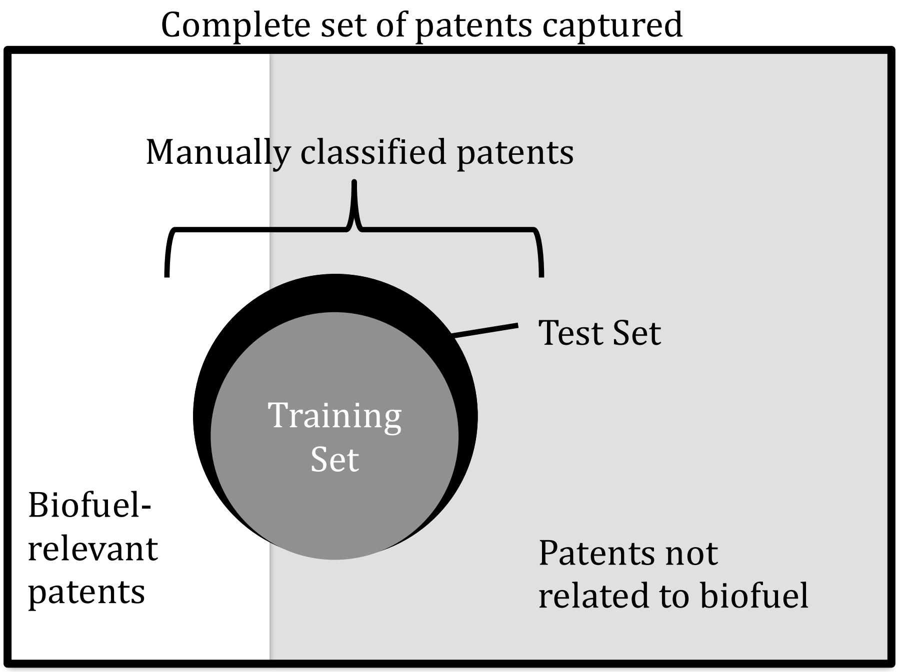

Figure 2.1. Graphical representation of the patent data captured and assessed to create and evaluate a biofuel classification model.

Costantini et al. (2013) created a unique biofuel-specific patent database built around keyword searches and patent classification codes from the WIPO Green Inventory. Their approach has been shown to substantially improve on the Green Inventory, a biofuel patent classification scheme where a set of relevant International Patent Classifications, believed to provide a good match to biofuel-relevant patents, is used for selecting biofuel patents. The NLP model, at the aggregate level, correctly classifies 84.7% of all biofuel patents. Further assessing the validated data reveals minimal error associated with the number of patents classified as “not applicable” that are actually biofuel patents. This error in classification accounts for 0.67% of all “NA” patents. Assuming this level of accuracy holds for the entire database of patents, the biofuel patent database effectively captured 58% to 80% of all available biofuel patents associated with the initial keyword searches. Based on the sample size and a confidence level of 99%, the classification approach put forward correctly identifies a patent as a biofuel-relevant patent 82% to 96% of the time. This is a substantial improvement in biofuel patent classification and categorization compared to either the green patent inventory (up to 14% accurate at selecting for biofuel-relevant patents) that makes use of the IPC system, or the BioPat database (up to 23% accurate at selecting for biofuel-relevant patents) as reported by Costantini et al. (2013).

Due to the overall dominance of non-biofuel related technologies, followed by 1st generation technologies in the random patent sample, I only use the classification results to assess patent trends for 1st generation and 2nd generation biofuel technologies, as well as for assessing trends associated with ethanol and biodiesel. These classifications could be expanded on by sampling more patents, or by using keyword filters or other first-pass classification methods to better select for certain types of biofuel-related patents that have been under-represented in the manually classified data.

The new international standard for patent classification, Cooperative Patent Classification, has since started to replace the IPC classification model and offers a number of improvements to IPC classification (Lasevoli, 2013). Because biofuel technologies have emerged out of several different sectors, it is likely difficult to capture all patents relevant to biofuel innovation with only a few CPC classes. The biofuel supply chain makes use of a myriad of technologies, that are all undergoing dynamic shifts (Yue, You, & Snyder, 2014). It is therefore important to not only capture biofuel production processes innovations, but also relevant feedstock innovations and other pathway innovations, such as enzyme development. The Y02E50 CPC classification, for instance, contains patents associated with technologies for the production of fuel of non-fossil origin, but is unlikely to capture relevant technologies for extracting saccharides from cellulose, as may be captured in the C12P19 class for the CPC. While the CPC system offers improvements to the IPC classification scheme, relying solely on CPC aggregation for patents is unlikely to provide the level of precision in classification compared to either the NLP method of classification presented here, or to the methods proposed by Costantini et al. (2013).

To provide additional assessment of CPC classification in relation to biofuel innovation system shifts, I have created an NLP-CPC concordance for biofuel. This is a table that provides a set of CPC classifications that capture many of the patents that my NLP method identified as being biofuel patents. This can serve to expand the CPC patent dataset to include patents from related fields that are not directly captured by the Y02E50 classification, but are captured through my NLP classification methodology. I have extracted the CPC classification code from each patent in my NLP-classified biofuel patent database, and created a frequency distribution for the occurrence of each CPC code. I then selected the top 20 most frequently occurring CPC classes for 1st generation, 2nd generation, ethanol, and biodiesel biofuel technologies. These classifications are provided in the appendix as a CPC-NLP patent classification lookup table. In Chapter 3, I refer to trends from this lookup table as CPC*. The CPC* classification approach contains many more patents than the other classification approaches, as it expands the scope of what is considered to be part of the biofuel technology innovation system, without removing the patents captured that are not associated with biofuel technologies.

Because biofuel technology could not exist without reliance on other sociotechnical systems, I have also collected additional patent data to elucidate trends in related sectors. 1st generation biofuels are often thought to originate from the agriculture industry while 2nd generation fuels originate from more-science based technologies, such as biotechnology and thermochemical processes (Suurs & Hekkert, 2009). Because it is believed that biofuels originate from two distinct sectors, I have collected patent data associated with the agriculture industry as well as patent data associated with biotechnology. Given the breadth of topics covered by these two sectors, I utilized traditional classification methods based on the IPC and CPC classification systems to establish patent count estimates for these sectors rather than relying on NLP approaches.

Biotechnology sector patenting activity was approximated by utilizing 30 USPTO patent subclasses as identified by Adelman and Deanglis (2007). I performed a count of all patents associated with these 30 subclasses from the USPTO from 1976 through 2013. For the agriculture sector, patenting activity was approximated through use of the CPC, where all patents from the A01 subclass ("Agriculture") were counted from 1976 through 2013.

Because biofuels are intended to displace traditional fossil resources in transportation, I have also collected patent data associated with the oil and gas sector. U.S. patent data was collected for the five largest oil and gas companies traded on the New York Stock Exchange in 2014 (BP, Chevron, Conoco, ExxonMobil, and Shell). Patents were collected by searching for all USPTO patents assigned to each of the five companies (9,300 patents). All patents assignee results were then reviewed manually, and any patents assigned to unrelated companies were removed.

2.2 TIS Analysis of Biofuels

In addition to using patent data to assess biofuel innovation, I turn to another data source for information: newspapers and trade journals. As previously established, innovation relies on the flow and use of information. Patents have traditionally been used as a proxy for knowledge stocks –capturing some of the information that inventors and entrepreneurs utilize which can ultimately lead to innovation (e.g. J. Wu & Shanley, 2009; Zucker, Darby, Furner, Liu, & Ma, 2007). However, innovation relies on a large network of actors and institutions working to generate, diffuse, and utilize information. This group of actors is unlikely to rely solely on patents for information, and so patents cannot be expected to capture the complete story of what's going on in a technology's innovation system (Moser, 2013). Instead, different agents rely on a number of different sources of information, of which patents may be a part (Acs, Anselin, & Varga, 2002). One other common source for information is textual media (either online or in print) (Tichenor, Donohue, & Olien, 1970).

Newspaper articles and trade journals are likely to report on important inventions and innovations as they gain social relevance (Acs et al., 2002). Often, articles may be speculative in nature, or may even be antagonistic toward a technology. As a technology becomes more relevant to society and to other niches, there is likely to be more reporting about that technology. If only a small number of informational sources are evaluated, there is the danger of over-representing the connectivity and size of the technology innovation system. This, in turn, could result in a "hype" bubble (Melton & Axsen, 2015). As more information is evaluated and considered, a more nuanced view of the innovation system should emerge, reducing the risk of hype cycles. In turn, hype can ultimately contribute to the development of stronger, more connected networks, by motivating actors and institutions to invest in the technology, further leading to the creation of new knowledge and knowledge flows.

Crowds have been shown to be efficient predictors of events – ranging from the probability of discovering the Higgs boson to movie box office sales and politics (Pennock, Lawrence, Giles, & Nielsen; Ray, 2006). Social media trends have similarly been shown to be good predictors for movie revenues, and the use of internet participants in games has even been shown to be effective at predicting low-energy configurations for proteins (Asur & Huberman, 2010; Savage, 2012).

Given the "wisdom of crowds" phenomenon, tracking technology as it is written about in the media may also provide a good proxy for technology innovation. Surowiecki (2005) writes that the "wisdom of crowds" is often good for making predictions about the likelihood of future outcomes, and when they will occur, as long as four factors hold: (1) Diversity: each person adds private information or bias, (2) Independence: people form their own opinions, (3): Decentralization: people draw their own specialized knowledge, and (4) Aggregation: a mechanism exists to turn private judgments into a collective action.

Given that newspapers and trade journals operate in different regions and cater to different reader demographics, their writers are likely to report on different technologies differently, and will gain information about these technologies through different networks. They may therefore choose not to report on technologies at all. This makes the information contained in newspapers both diverse and decentralized. Because the content being published is ultimately left to the editor, information pertaining to the technologies featured and how often a specific technology is featured is likely to be established independently, and take into account the likely importance to the readership of that specific source. As such, aggregation of newspaper data across a large enough set of media sources is likely to exhibit "wisdom of crowds" behaviors for capturing innovation system trends, and therefore reduce the risk of capturing "hype," which may happen if only a small number of newspaper sources are utilized (Melton & Axsen, 2015).

For instance, as electric vehicles deployment increases across the United States, we would initially expect to see articles and information associated with EVs to show up in news sources relevant to the areas that receive electric vehicles first, or areas that are more interested in EV technology. As time goes on, other local newspapers and news stories are likely to write about EVs, or to more frequently mention of the technology. If some regional news sources do not think electric vehicles are relevant, they may choose not to write about or discuss the technology. As the technology becomes increasingly more indispensible to society, there should be more articles discussing that technology overtime. Take gasoline and oil, for example, where prices and price implications and changes are often discussed across a wide variety of news sources, with entire sections dedicated to the discussion of oil and gasoline market trends.

If a large enough number of informational sources are drawn from, a more accurate representation of the technology can be expected – more important and socially relevant technologies are likely to receive more press (be it from pricing reports, or deep-dives into technological opportunities and uses). The reporting, or lack of reporting, that takes place across a large set of information sources is likely to result in different trends than if only larger, national sources of information are drawn from to make conclusions, such as the New York Times or Wall Street Journal.

In addition to providing a wisdom-of-crowd approach for assessing technology, articles allow for more nuanced categorization and classification of innovation trends than is possible with patents, similar to how patents provide further detail and disaggregation on the innovation process than simply looking at R&D investments. Negative press about EVs, or highlights of EV deployment failure, for instance, can be captured. This could include instances where batteries explode, or where someone reports on how their EV ran out of electricity two-miles from their home. This negative press could ultimately detract from the advancement of the technology.

Across different regions, and over time, articles will report on technologies in a variety of different ways. Events that are important receive wider attention and increased replication. As industries mature, reporting is likely to increase, and new information sources (such as trade journals) may emerge. Additionally, the information contained in newspapers and trade journals can provide a more immediate snapshot of the innovation system than is possible with patents. Despite the seeming appeal of newspapers as a means to assess innovation, there are still drawbacks. One of the biggest drawbacks of newspaper data is that access to the data is often limited. There are a myriad of different news sources across regions; not all of them are in the same language, and not all of them have the same publishing reach. Capturing articles from a large variety of sources can be a difficult task. Choosing which sources to draw from while ignoring other sources may bias the results. Furthermore, not all newspapers maintain long-term digital archives that are publicly accessible. Without capturing a large enough set of diverse newspaper sources, it is likely that the data will be biased, and not exhibit “wisdom of the crowd” behaviors.

In addition to the boundary condition and data source concerns, there is also one other substantial concern: classification and use of the data. Once newspaper data is collected, it can be difficult to use the data in a quantitative or meaningful fashion. One individual might be able to read hundreds, or even thousands of newspaper articles related to a given subject, but it would be impossible for that individual to read all newspaper articles that could be captured. There is therefore a tradeoff between establishing a large enough newspaper dataset to work with so that trends remain unbiased, and in being able to read and gain insight from the data.

To relate newspaper articles to technology innovation, I rely on an event analysis approach to article classification. That is, I assume that each article represents an event, or a specific occurrence that is relevant to the technology innovation system. For instance, if legislation were passed to provide a tax incentive for biofuels, this would mark an event that occurred and is being written about to inform the public. Alternatively, if a long-form informational piece about biofuels is published, the article itself may represent the event, or may reflect a non-obvious event that triggered the writing or publication in the first place (e.g. a biofuel production facility in a nearby county went bankrupt or was involved in a scandal). More important events are likely to receive more media attention across a wider variety of newspaper sources. In this fashion, articles that effectively duplicate the same new story help to internalize the magnitude of the effect of that event.

Event analysis, however, is a non-trivial, time-intensive process that is not without faults and limitations. Event analysis has been used in a number of different innovation studies (e.g. Tigabu, Berkhout, & van Beukering, 2015; Van de Ven & Poole, 1990), and has previously been used to investigate technology innovation systems for European biofuels (Negro, Hekkert, & Smits, 2007; Suurs & Hekkert, 2009).

The event analysis work by Suurs and Hekkert (2009) uses a number of journal articles associated with specific biofuel topics. Their team sorted through those articles to find a small subset of articles that were then selected to inform a technology innovation narrative. Similarly, work by Negro et al. (2007) uses a small subset of articles and media sources to construct an event narrative. Inherent to these event analysis approaches is general content analysis, where the content of specific written media is read, analyzed, and coded so that details can quickly and readily be accessed through a database. This process can be very detailed, listing frequency distributions of the words used in a specific article, for instance, or can be more general, like specifying the subject of the article with a single keyword.

Content analysis is not only time-intensive, but also subjective in how articles are grouped and coded. Furthermore, there is selection bias associated with the media sources utilized as only a small subset of sources can be read in full and coded (Grimmer & Stewart, 2013; Hopkins & King, 2010)

Although TIS event analysis, relying on content analysis, has its limitations, new approaches have made headway in overcoming some of these limitations. Specifically, content analysis methodologies employed by Sengers, Raven, and Van Venrooij (2010) substantially expand the article datasets usable for TIS analysis through use of a computational tool kit for content analysis, T-lab. T-lab allows articles to be clustered by word choice similarities and placed into specific groups. By breaking a large set of articles into groups, temporal trends across groups can be evaluated, leading to additional insights for technology innovation and shifts in the TIS over time without needing to manually read and code each article. Approaches like this can help to remove some of the selection bias that is inherent to content analysis. While Sengers et al. (2010) track specific shifts in computationally established clusters, it is unclear how each cluster relates to overall TIS outcomes, or if this approach is best-suited for TIS analysis.

Rather than relying on pre-built content analysis software that is limited in scope to word clustering to derive technology groupings and trends, I combine content analysis techniques alongside computational linguistics to further improve upon traditional event analysis. Following the approach I used for classifying patents, I utilize a supervised learning method to train a myriad of custom machine-learning models that make use of NLP to better classify and sort articles into specific TIS functions. The 7 TIS functions were developed for the purpose of better classifying and grouping the types of actions that can be taken to promote the formation of structural elements of the innovation process. Small-business loans, for instance, could be used to increase the number of entrepreneurial firms in the market. Alternatively, research grants could be used to increase basic knowledge about the technology. These are direct actions that can fit within the 7 different TIS functions and can be used to promote a specific kind of network formation.

For this analysis, I created 15 binary models to classify each document by technology type (e.g. ethanol, biodiesel, etc.), technology generation (1st or 2nd generation), direction of support (blocking/hindering or supporting/promoting), and innovation function (Table 2.3). All texts I collected were tested against each of the 15 models to provide unique classification results.

The text corpus I use is made up of newspaper articles and metadata associated with articles from the LexisNexis Academic newspaper database. LexisNexis tracks over 6000 English-language news sources, each containing a large number of articles and original content. To make the data collection process feasible, I only chose to analyze biofuel texts from 1995 through 2013 – the complete collection of texts analyzed was approximately 973,000.

The initial article repository was established by conducting keyword searches on LexisNexis with the following 6 keywords, relevant to biofuel technology: Biofuel, Biodiesel, Drop-in Fuel, Ethanol, Renewable Diesel, and Biogas. Given the broad nature of the search terms, I believe that the articles captured are likely to contain most print-media articles concerned with biofuels for the available sources.

I used the built-in LexisNexis algorithms to find and remove duplicate articles from within keyword searches, and utilized title and date identifiers to remove duplicate articles across keywords (if an article showed up in both Ethanol and Biodiesel keyword searched, it would not be counted twice). Removing all duplicate articles created a newspaper repository of approximately 634,000 biofuel-related articles.

I performed the initial content analysis across a small subset of articles; 1001 articles from the newspaper data repository were randomly selected and read in full. To better elucidate meaningful trends associated with the technology innovation system, I read and classified articles according to TIS functions (Table 2.3). I also classified articles into technology types, technology generation, and the direction of support (supporting or blocking).

I used only two technology generations for this classification analysis: 1st generation and 2nd generation biofuel technologies. Articles discussing technologies that make direct use of conventional food crops were viewed as 1st generation technologies, while technologies that use non-food or non-conventional resources were classified as 2nd generation technologies. 2nd generation technologies include technologies like cellulosic ethanol production from corn stover using biochemical or thermochemical pathways, and biodiesel production from algae.

For each of the 7 TIS functions, the supporting classification was defined based on the tone and the direction of functions exhibited in the article. If, for instance, the article mentioned a government-issued tax incentive for ethanol, that article was classified as supporting the technology; if, however, the article mentioned removal of a tax incentive, or discussed an opinion that the tax incentive should be removed, that article was considered a blocking event. Manual classification allowed for each article to capture more than 1 TIS function and several different technologies. Table 2.3 shows the general criteria I used for manually classifying articles into functions. These criteria were established as guidelines for making decisions on classification. Other experts in the field may develop a similar, or more comprehensive set, which may result in slightly different classification results.

Table 2.3

Criteria for article classification

Function	Characteristics of Articles
Resource mobilization	Government subsidies, research grants, venture capital, IPOs, direct investment of capital from third parties
Market formation	Government policy incentives, long-term contracts for product
Legitimation	Sign of product usage, product discussion, showing of support, success or failure stories, expectation or projections
Knowledge development and diffusion	Discussion of R&D efforts, educational events, patent implications or filed applications, “Basic Research” or Knowledge sharing, pilot facilities
Influence on the direction of search	Government policy incentives, highlights of positive or negative commercial outcomes
Entrepreneurial experimentation	Discussion of joint-ventures, commercialization activity, business mergers and acquisitions, launching a new product or production facility
Development of positive external economies	Involvement of seemingly different sectors, usage of product for different applications

I built 15 different NLP models, which enabled me to analyze all 634,000 articles, attaching 15 different classifications to each article. The classification models I created made use of several article characteristics and parameters. First, each article was split into specific words; word tokens in each article were defined as items starting with a letter, followed by any number of ASCII digits, numbers, percent expressions, money, and white spaces. I then stripped and ignored the white spaces around each word token. The length of each article was established based on binned sets of tokens: 500, 1500, 4500, and 13500 word tokens. Word pairs were also utilized, but word triplets were not found to improve classification results. This resulted in a set of unique characteristics and words to describe each document analyzed. The Stanford NLP Classifier than utilizes statistical models and algorithms to determine the set of features that most consistently predict the correct article classification. The trained classification models used roughly 24,000 different textual features in the decision making process for whether or not an article belongs to a specific technology class.

Because I created 15 different classification models, each article can be classified as fitting one or more different technology, and as supporting or blocking one or more different TIS functions. Of the 1001 articles that I manually classified, I utilized the same random selection of 770 articles to train each of the 15 different classification models. The remaining 231 articles were later utilized to perform model validation and to assess model accuracy, precision, and classification error. Aggregate details for article classification based on manual classification and the classification model output is provided in Table 2.4.

Table 2.4

The number of articles classified into each category manually, and automatically using NLP algorithms

Classification Model	Manually Classified count	% of manually classified articles	Automatically classified count	% of all articles
1G	852	85%	620287	98%
2G	197	20%	58408	9%
Biodiesel	389	39%	231295	36%
Biogas	95	9%	39560	6%
Blocking	192	19%	74028	12%
Development of Positive External Economies	292	29%	55172	9%
Drop-in	37	4%	455	0%
Entrepreneurial Experimentation	270	27%	144391	23%
Ethanol	511	51%	336331	53%
Influence the Direction of Search	178	18%	63200	10%
Knowledge development and diffusion	116	12%	22606	4%
Legitimation	607	61%	434230	68%
Market Formation	143	14%	61748	10%

Resource Mobilization	146	15%	60141	9%
supporting	776	78%	611353	96%

As evident by Table 2.4, there is a substantial difference for the percent of articles classified into each class for the manually classified data subset and the overall article database (to which the classification models were applied). This difference can likely be attributed to the modeling error for each classification model, or may indicate that the initially classified articles were not selected in a truly random fashion. At a 95% confidence level, only article counts attributed to the ethanol and biodiesel classification models show no significant deviation from the manually classified counts. To better ascertain model quality, I have validated the models by using the remaining 231 articles that were manually classified and not used to build each model.

Figure 2.2 shows that model accuracy is above 60% for all models (the number of biofuel articles and non-biofuel articles correctly classified). The recall value indicates the percent of the in-class signal that is accurately captured through the classification model; this value shows the accuracy of the model for only the correctly identified positive results, and does not include the correctly identified negative results. In other words, the recall value for 1G biofuels represents the number of 1st generation biofuel articles that were correctly classified as 1st generation biofuel articles, and does not consider the number of articles that were correctly classified as being unrelated to 1st generation biofuels.

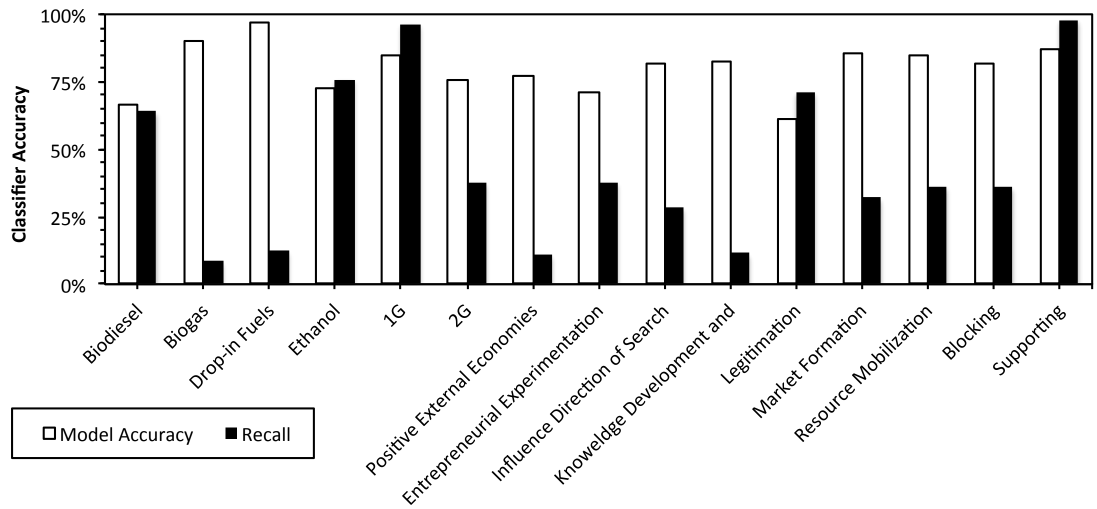

Figure 2.2. Accuracy of the supervised learning methodology employed to create 15 different binary classification models.

When looking at recall values, it appears that some of the classification models perform substantially better than others. As is the case with machine-learning algorithms, classification accuracy is strongly correlated with the size of the training set used, in which larger training sets can often improve classification results (Foody & Mathur, 2004; Hopkins & King, 2010; Zhuang, Engel, Lozano-Garcia, FernÁNdez, & Johannsen, 1994).

For the models used to classify articles into the 7 innovation functions, the average model accuracy is 78%, with an average recall value of 0.33 and a mean average precision of 0.60 (Figure 2.3 graphically represents recall and precision). These values are substantially higher or on par with similar in-class classification methods utilized. For instance, default patent classifications often capture a low number of patents relevant to the technologies being studied (Costantini et al., 2013; Herbert, Szarvas, & Gurevych, 2010; C.-H. Wu, Ken, & Huang, 2010). Even the BioPat database, a database established by

Costantini et al. (2013) to improve patent classification for biofuels, shows a recall value of only 0.23.

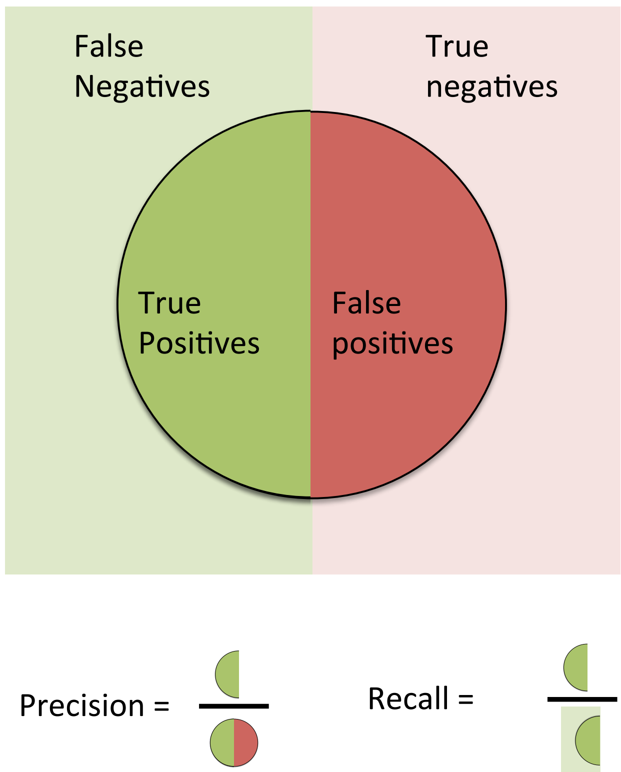

Figure 2.3. Graphical representation of precision and recall for information retrieval.

Usage restrictions prevent me from making article data from the LexisNexis article database publicly available. Instead, I have publicly released classification identifiers alongside article titles and sources when possible (Kessler, 2015a).

2.3 TIS Analysis of EVs

The final portion of this dissertation is devoted to assessing an emerging transportation technology. Specifically, I choose to focus on electric vehicles.

Given that the innovation narrative for electric vehicles is still developing, I met with numerous stakeholders to gain insight into what has occurred. In turn, I compare the narrative that emerged through stakeholder discussions with the innovation narrative that can be revealed by the textual analysis methodologies used in the previous sections.

To better ascertain how the innovation process is unfolding for electric vehicles, I met with a number of important and relevant actors in the EV landscape for three different states. These actors come from a variety of different organizations, including state and local governments, non-profit groups, car dealerships, utilities, infrastructure companies, and research institutions. This research was done as part of a greater study (results forthcoming) supported by The National Center for Sustainable Transportation and the Zero Emission Market Acceleration Partnership (ZE MAP) at the University of California, Davis. I communicated with representatives from three main states: Washington, Colorado, and Georgia.

Utilizing contacts from ZE MAP, an email was sent out to point people in Washington, Colorado, and Georgia. The email requested that these contacts provide a set of actors that they thought were relevant for promoting electric vehicle uptake within the state. An example of the form used to collect this data is shown in Figure 2.4.

Entry #	Stakeholder category	First Name	Last Name	Email Address	Affilitation	Strength of contact	Other Notes
Example		John	Smith	example@example.com	EV Motor Corp	5 - Know this individual well or have worked with them	This is an example field
1	**State government**						
2	**Local/regional government**						
3	**Car company**						
4	**Electric utility**						
5	**Non-profit organization**						
6	**Research institution**						
7	**New car dealership**						
8	**Infrastructure company**						
9	**Technology company**						
10	**Other**						

Figure 2.4. Form used to collect possible stakeholders for in-person meetings

From these emails, I obtained a set of contacts to schedule for in-person meetings. Each contact was emailed, and a set of possible times and dates were provided for them to meet. Response rates were greater than 80% for each state.

To better construct a technology innovation narrative, Gustavo Collantes, the lead ZE MAP researcher, and myself developed a protocol to guide the discussion. Discussion was structured around the seven functions of the Technology Innovation System framework. Careful consideration was made for developing each question to better align the discussion with the core attributes of TIS functions. Meeting participants were asked to provide information related to electric vehicle uptake in their state, which included both fully battery-electric and extended range hybrids (Plugin Electric Vehicles - PEVs).

To assess **Resource Mobilization**, information was gathered that pertained to the following questions:

1. Could you share with us information about amounts and allocation of funding related to PEVs?

2. Are more resources needed?

3. What are the main sources of uncertainty for whether resources will be allocated now and in the future?

4. What role is the Governor's Office playing in PEV discussions in your state?

To assess **Market Formation**, information was gathered that pertained to the following questions:

1. Are there local incentives in place to help users and fleets adopt PEVs?

2. What kind of support/incentive is there for the deployment or use of charging equipment?

3. Have there been marketing campaigns developed to generate public interest in PEVs?

To assess **Legitimation**, information was gathered that pertained to the following questions:

1. How informed is the general public in your state about PEVs?

2. What images do people associate with PEVs? Are these positive or negative images?

3. Who are the primary market segments that purchase PEVs in your state? Why?

4. Are there important stakeholders that are antagonistic toward PEVs?

5. Is there market demand, and does local government or advocacy groups support PEV adoption?

To assess **Knowledge Development and Diffusion**, information was gathered that pertained to the following questions:

1. What are the more critical factors that drive/deter the market adoption of PEVs in your state?

2. How do stakeholders in your state learn about the fact that these factors are critical?

3. Do most stakeholders agree on the importance of these factors?

To assess **Influence on the Direction of Search**, information was gathered that pertained to the following questions:

1. How do existing state/local actions influence activities to support plug-in vehicle adoption?

2. What laws, regulations, competing technologies, or organizations prevent PEVs from getting more attention from stakeholders?

3. How has media affected PEV adoption in your state?

4. Is there a PEV action plan in place?

5. Are there automaker or fleet EV requirements?

6. Have there been clear statements or expression in support for EVs from the political leadership?

To assess **Entrepreneurial Experimentation**, information was gathered that pertained to the following questions:

1. Are there companies in your state that supply parts, services or technology for the production of PEVs or charging infrastructure?

2. Are there exceptional PEV champions in your state?

3. Have there been innovative ideas that were proposed and tested in your state to support PEV markets?

4. Do you know of companies or organizations that are developing and/or testing new technologies or services?

To assess **Development of Positive External Economies**, information was gathered that pertained to the following question:

1. What external, economic benefits do you think EVs can offer? This includes things such as fostering jobs, environmental benefits, or aiding/supporting alternative industries

Over a 1-month period, 31 individuals across 3 different states participated in meetings to assess EV Innovation. Meetings took place during a 45-minute period, in which questions were asked in accordance with the established protocol. Meeting notes were summarized and retained. These notes have been used to inform the EV technology innovation narrative for each state. Responses from the individuals have been kept anonymous to better protect participants. In addition to utilizing these notes as an important dataset for my dissertation, the responses are also being incorporated into work by the National Center for Sustainable Transportation and ZE MAP (not-yet published).

Alongside stakeholder discussions, I have also used data from ZE MAP associated with electric vehicle deployment for each state at a monthly interval from January 2010 through November 2014. This data contains vehicle make, model and registration information, so that monthly uptake associated with electric vehicles can be determined. The vehicle registration dataset consists of 24 states: Arizona, California, Colorado, Connecticut, Florida, Georgia, Illinois, Indiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, New Jersey, New Mexico, New York, North Carolina, Oregon, Rhode Island, Tennessee, Texas, Vermont, Virginia, and Washington. Additional monthly data for each of these states has been obtained from ZE MAP which includes rural and urban vehicle miles traveled, and average residential electricity prices. Gasoline prices were obtained from the Energy Information Administration for all states in 2010 (Energy Information Administration, 2015). The average monthly price for each month in 2010 was used to model monthly gasoline prices in each state over time.

Newspaper data associated with electric vehicles was collected to assess innovation system trends as portrayed by print media. The text corpus I use is made up of newspaper articles and metadata associated with articles from the LexisNexis Academic newspaper database. I have collected data from 1994 through 2014 – the complete collection of electric vehicle articles analyzed was approximately 134,000.

The initial article repository was established by conducting a keyword search on LexisNexis using the following keyword: Electric Vehicle. I used the built-in LexisNexis algorithms to find and remove duplicate articles and also utilized title and date identifiers to remove any additional duplicate articles. Removing all duplicate articles created a newspaper repository of approximately 133,500 EV-related articles.

Again, I utilized machine-learning algorithms and a small subset of manually classified articles to provide insight into the technology innovation system based on the collected newspaper data. I selected a small subset of random articles from the electric vehicle database. These 1000 articles were manually classified into the 7 technology innovation system functions. The same criteria for classification as shown in Table 2.3 were utilized for manually classifying these articles.

To further validate this manual classification methodology, I trained another student to identify and classify articles into corresponding innovation system functions. An initial 45-minute training session took place where the student was exposed to the fundamental aspects of the TIS theory, and introduced to the content analysis system that I designed. The student and I read articles, and specified TIS classifications, discussing overall categorization until classification consensus was achieved. Training was believed to be adequate once consensus had been achieved for five articles in a row prior to discussion. Feedback from this student on the classification process is provided below:

> *The categories overall seemed clearly defined. Most of the difficulty in making the distinction between categories was gray area in the extent to which an event warranted inclusion in one or more categories. For example, an a topic may have elements of knowledge development and diffusion, but to a very weak degree. Additionally, some articles mention plug-in vehicles to a varying extent, so it is up to the evaluator to judge between relevant and not applicable to the plug-in category.*

While manual classification of articles is not without problem, I was able to demonstrate consistency in teaching and applying this process. This is indicative of the possibility to further extend this methodology to other fields of interest, and to assess other technologies.

From the 1000 manually classified articles, 700 were utilized to create a classification training set. The remaining 300 articles were utilized to validate the classification models. 10 different binary classification models were created: one for each of the 7 different innovation functions, one for articles that support the functions (those articles that are positive about the technology, likely promoting electric vehicles), and one for articles that block functional fulfillment (articles that are negative about the technology, or report on negative outcomes, such as battery failures), and one to classify if an article did not relate to the EV Innovation System, and was therefore inapplicable. Classification validation results from this are shown below.

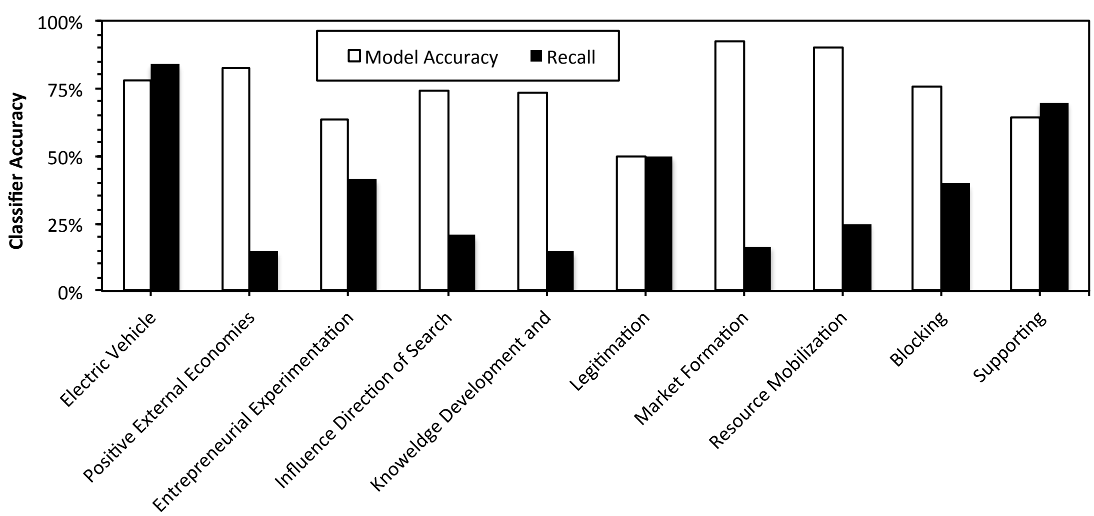

Figure 2.5. Accuracy of the supervised learning methodology employed to create 10 different binary classification models.

As indicated in Figure 2.5, model accuracy is above 50% for all models. The number of false negatives and false positives reported by each model, however, skews model accuracy. The recall value indicates the percent of the in-class signal that is accurately captured through the classification model. For instance, this value shows the number of articles manually classified as knowledge development and diffusion that were then automatically classified as knowledge development and diffusion. It does not consider the articles that were correctly identified as not relating to knowledge development and diffusion. The mean average precision is 0.55.

To further test the defining characteristic of the 7 innovation functions as established in the TIS literature, I conducted robustness tests using 5 different TIS classification models on a different subset of 324 manually classified electric vehicle articles.

Model 1 (Biofuel Naïve): This is a Naïve model, or a simplistic model with limited complexity that acts as a baseline guess for what will occur without using additional data. The general ratios of TIS functions in a small subset of biofuel articles (700) were used to extrapolate likely trends for the EV TIS

Model 2 (EV Naïve): This is a Naïve model, or a simplistic model with limited complexity that acts as a baseline guess for what will occur without using additional data. The general ratios of TIS functions in a small subset of EV articles (300) were used to extrapolate likely trends to the EV TIS

Model 3 (Biofuel ML): The Biofuel machine-learning (ML) model Without further modifications or re-training, the biofuel classification model (discussed previously) was utilized to classify EV articles into TIS functions.

Model 4: A machine-learning model created using the remaining EV articles not in the 324 EV article test set.

Model 5: A pooled model, where a training set of 676 EV articles and the complete set of manually classified biofuel articles were combined to create a pooled machine-learning training set, made up of two markedly different technologies.

Figure 2.6 shows how well each model predicted the number of articles associated with a given TIS function in the article subset. If the bar falls above the line, the model over-predicted the number of TIS functions of a given type in the actual dataset. If the bar falls below the line, the model under-predicted the number of TIS functions of a given type in the actual dataset.

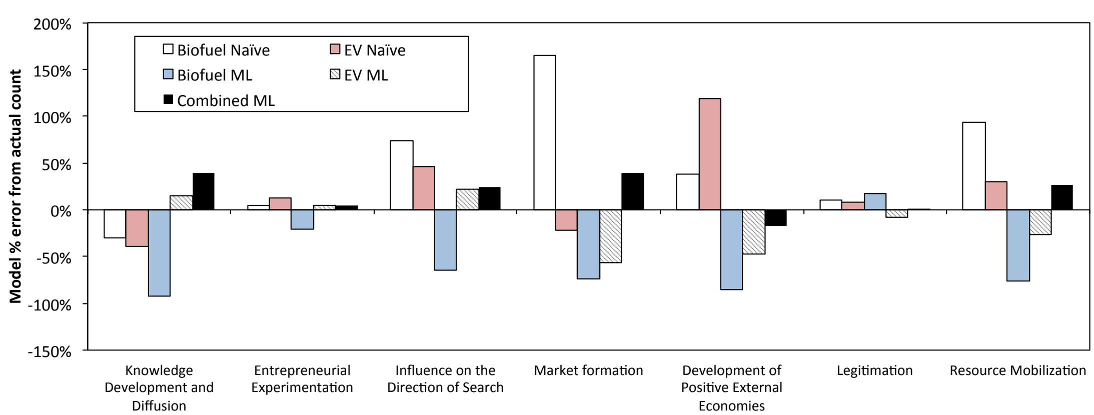

Figure 2.6. Percent difference from actual article counts – the machine learning models show better fit to actual article counts compared to the Naïve models

As evident in Figure 2.6, the Naïve Biofuel model substantially over-estimates the number of articles associated with a given innovation function for all but the Knowledge Development and Diffusion function. This model is useful for directly comparing how the biofuel TIS has been supported relative to how the EV TIS has been supported – the larger the magnitude of the bar, the greater disparity there is between the two innovation systems for a given function. Both the Combined ML model and the EV ML model substantially out-perform the naïve forecasts.

The machine-learning model for Biofuel classification does not improve EV classification results compared to the Naïve EV model (blue and red bars respectively). The magnitude of classification error for both the Biofuel ML model and the Naïve EV model, however, is similar for many of the functions being assessed, although the Biofuel ML model consistently underestimates the number of EV-related articles associated with each innovation function.

The Biofuel ML model therefore provides conservative estimates of TIS functional support for EVs compared to regularly overestimating the functional support for EVs as is done by the Biofuel Naïve model. This suggests that a machine-learning approach to classifying articles could be used to provide a first-pass, conservative estimate for how well supported a technology's innovation system is. Increasing the number of articles and technologies utilized in creating a TIS-classification model could help to eliminate some of this error, as indicated by the Combined ML model.

To further assess the objectivity of classification models in predicting TIS functional fulfillment, I create a "pooled," or combined classification model. For this model, I utilize 676 manually classified articles from the electric vehicle newspaper dataset, and combine this with 1000 manually classified articles from the biofuel newspaper dataset. This combined set of 1676 articles is utilized as the training set, which is then validated against the remaining 324 articles pertaining to electric vehicles. Classification accuracy from this approach is shown below.

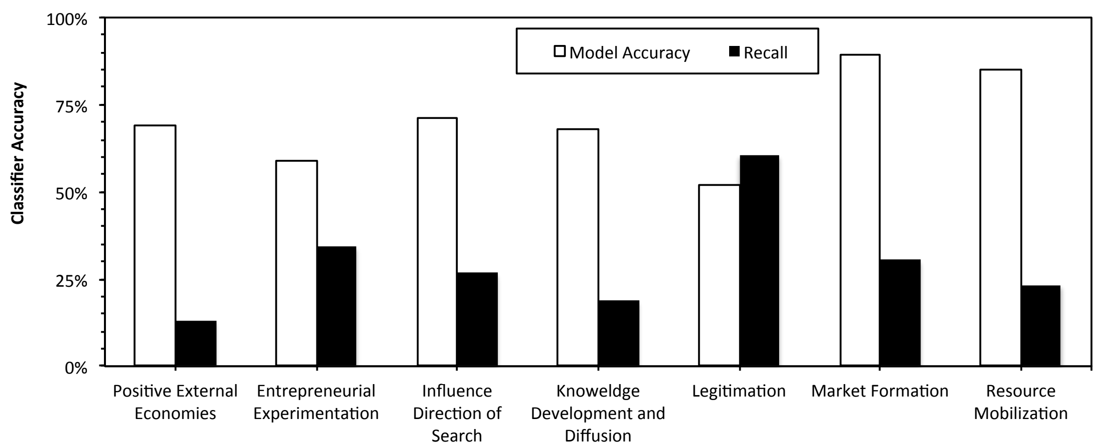

Figure 2.7. Accuracy of the supervised learning methodology utilizing two distinct technology classification subsets

Based on the results of the pooled classification model, in addition to the model robustness tests shown in Figure 2.6, I am led to believe that given the size of the dataset and the independent nature of the two topics, customized innovation models are likely to yield better results than pooled models. Nonetheless, classification accuracy is still reasonably high across the pooled dataset; all models are at least 50% accurate, and recall values are on par with classification accuracy of other innovation data sources, such as patents (Costantini et al., 2013). Given these results, and that the mean average model precision is 0.52, the pooled NLP model could likely be used as a first-pass option for estimating technology innovation fulfillment for other technologies. All ML classification models used have been made publicly available (Kessler, 2015b).

The ML EV model has been used to classify the remaining articles in the electric vehicle article database. Once classified, articles were filtered using LexisNexis metadata to correspond with state geographies. Analysis and results for electric vehicle innovation is further discussed in Chapter 5.

Chapter 3: Assessing Patents as a Proxy for U.S. Biofuel Innovation

In this chapter, I explore the use of natural language processing (NLP) as a means to improve classification accuracy and precision for biofuel patents. This dataset is later used in Chapter 4 to compare the validity of using a different innovation proxy. Trends and patterns associated with patenting are highly dependent on how patents are classified and grouped together. I find that U.S. biofuel patenting activity closely corresponds to the sociotechnical shifts described in the biofuel innovation narrative presented below.

Using NLP and machine-learning algorithms, I am able to expand patent classification capabilities, by capturing a greater selection of patents, and by classifying relevant biofuel patents more precisely than could be done with keyword searches alone. I find a better match for patent activity with the U.S. biofuel innovation narrative when using the NLP-derived patent dataset compared to standard patent classification schemes.

Results indicate that after the initial establishment of the biofuel industry, there were two surges in biofuel innovation: 1995-2000, characterized by heavy patenting by 1st generation (food-based) biofuel firms; and 2005-2010, characterized by a second surge of innovation by those same large firms, complemented by a large number of biotechnology firms producing a relatively small number of 2nd generation biofuel patents. This analysis corroborates the idea that the first surge in biofuel innovation was linked to innovation in genetically engineered food crops, while the second surge of biofuel innovation was driven by policies mandating and incentivizing biofuels, with the biotechnology industry cautiously experimenting with non-food-based biofuels.

3.1 The U.S. Biofuel Innovation Narrative

This biofuel innovation narrative was developed through a review of the relevant literature, in addition to discussion with experts. Experts have reviewed the narrative, and feedback from this review process has been incorporated.

The 1970s oil crisis spurred alternative energy research in the United States, rekindling interest in the use of biofuel for transportation (United States Congress, 1979). Biofuel production and innovation in the U.S. traces its origin to the energy policies enacted in the late 1970s. Substantial support for biofuel first emerged in 1978 in the form of a tax-exemption of $0.40 (1978 USD) per gallon of ethanol. This tax measure remained in place, at various levels, for several decades, expiring only at the end of 2011 (Eggert & Greaker, 2014; Tyner, 2008). During this same time period a $0.40 (1980 USD) per gallon import tariff was also imposed against ethanol imports, serving to encourage the development of a domestic ethanol industry. From 1980 through 2000 the ethanol tax credit directed between $8.6 billion and $12.9 billion dollars ($2006) in support for ethanol (Koplow, 2006).

This generous ethanol subsidy in conjunction with high crude oil prices effectively established an ethanol industry in the U.S. (Tyner, 2008). With oil prices remaining high, the fixed ethanol tax credit facilitated rapid growth of ethanol. The U.S. biofuel industry expanded production from almost nothing at the start of the 1980s to about 1.3 billion gallons of ethanol by 1993. Modest support for this expansion also came in the form of The Intermodal Surface Transportation Efficiency Act of 1991 which provided states with transportation project funding for increased use of ethanol (Kelly & Brannon, 1996), and the Energy Policy Act of 1992 which encouraged governments at all levels to purchase alternative fuel vehicles for their fleets so that ethanol could be consumed at higher quantities (Corts, 2010; Winebrake & Farrell, 1997). Although the ethanol tax incentive remained in place, production volumes stagnated from 1993 through 1999, as ethanol demand became saturated.

At the same time that growth in ethanol production slowed, other socioeconomic shifts were underway. The agriculture sector went through a technological transition when a subset of seed companies started using genetic engineering technologies to improve specific traits of seeds. In 1996, the first commercial genetically engineered (GE) crops became available, offering higher crop yields and increased resilience to pests and herbicides (Dill, 2005; Fernandez-Cornejo, 2009). GE crops quickly became dominant in the United States (Fernandez-Cornejo, Wechsler, Livingston, & Mitchell, 2014). The abundance of new crop technologies alongside continuous federal support for ethanol created a positive feedback loop (Fausti, 2015). U.S. agriculture and energy policy choices merged to incentivize U.S. corn production. Further advancements in GE seed and ethanol production technologies facilitated efficiency gains in biofuel production (Fausti, 2015).

New demand for ethanol materialized in 1999 when California, followed by other U.S. states, adopted legislation to ban the use of methyl tert-butyl ether (MTBE), an anti-knocking fuel additive blended into gasoline (U.S. Environmental Protection Agency, 2004). Alongside the substantial subsidies for ethanol production, increased corn crop yields, and the need to replace MTBE as an oxygenate, the ethanol industry could one again grow production volumes, increasing production to 3.4 billion gallons of ethanol in 2004.

The U.S. Biodiesel industry can similarly trace its origin to government intervention in the 1990s. The Energy Policy Act of 1992, once modified by the Department of Energy, recognized biodiesel as an alternative fuel capable of complying with alternative-fueled vehicle mandates for government and state motor fleets (Koplow, 2006). Additional support for biodiesel came with the establishment of the National SoyDiesel Development Board in 1992, which later became the National Biodiesel Board (Howell, 1997). In 1993 dozens of demonstrations of biodiesel began, and in 1996 the first two major biodiesel companies started commercial scale production.

Additional support for the biodiesel industry came in 1998 with the Conservation Reauthorization Act, a law that amended the Energy Policy Act of 1992 to include biodiesel fuel-use credits (Koplow, 2006). This added cash support for the Bioenergy Program (M. Carriquiry, 2007). Annual capacity for biodiesel production increased during this time period from only 0.5 million gallons in 1999 to 20 million gallons in 2003 (Koplow, 2006). The American Jobs Creation Act of 2004 provided the first tax subsidies targeted directly at biodiesel, a model that closely followed the ethanol excise tax credit implementation. Biodiesel derived from virgin vegetable oils or animal fats earned a credit of $1/gallon, while biodiesel from waste oils earned $0.50/gallon (Koplow, 2006). With implementation of the tax credit, biodiesel production increased from 20 million gallons in 2003 to 112 million gallons by 2005 (M. Carriquiry, 2007).

Starting in 2005, the Energy Policy Act established the Renewable Fuel Standard (RFS) in addition to a small biodiesel producer tax credit. The RFS called for production of 4 billion gallons of renewable fuels in 2006, rising to 7.5 billion by 2012 (Koplow, 2006). The RFS was the first policy to provide direct support for 2nd generation and advanced biofuel, biofuel derived from non-food or non-conventional resources, by allowing cellulosic-derived ethanol to count as 2.5 times that of corn-based ethanol for complying with the standard (Solomon, Barnes, & Halvorsen, 2007).

In 2007 the RFS was renewed and expanded (RFS2), greatly increasing volume requirements for 1st generation biofuels, but also creating specific requirements for the phase-in of 2nd generation biofuel technologies (Tyner, 2012). The RFS2 required the use of 36 billion gallons of renewable fuel by 2022. Within the nested volume structure of the policy, only 15 billion gallons were allowed to come from conventional, starch-based ethanol, while the bulk of the mandate required advanced or cellulosic fuels for compliance. With the RFS2 in place, ethanol consumption continued to increase, growing to 13 billion gallons of fuel by 2010, or just over 9% of the U.S. gasoline market by volume (U.S. Energy Information Administration, 2015a). The biodiesel market remained much smaller than the market for ethanol, growing to 250 million gallons by 2006 (M. Carriquiry, 2007). 2014 Production of biodiesel was just under 1.3 billion gallons of fuel (U.S. Energy Information Administration, 2015b). Despite the rapid growth of biofuel following implementation of the RFS and RFS2, concerns over biofuel were soon raised due to uncertainty regarding the environmental impacts of biofuel alongside the global food crisis of 2007 and 2008 (Janda, Kristoufek, & Zilberman, 2012; Searchinger et al., 2008).

1st generation biofuel concerns have escalated, and an increased sense of urgency has been developed to move 2nd generation biofuels into commercial production (M. A. Carriquiry, Du, & Timilsina, 2011). In general, it is believed that 2nd generation biofuels may increase the sustainability of biofuel production (Balan, Chiaramonti, & Kumar, 2013). Although the RFS2 mandated the production of cellulosic biofuels, originally requiring 100 million gallons of fuel in 2010 before ramping up to 1 billion gallons by 2013, it took until

2012 for any cellulosic biofuels to be delivered. In 2012, only 20,000 gallons of cellulosic ethanol were produced, a fraction of the original expectation (Schnepf & Yacobucci, 2012). Current volumes of cellulosic and advanced biofuels remain far below original expectations, with no strong indication that the industry will mature to the levels desired in the near-term (Fulton, Morrison, Parker, Witcover, & Sperling, 2014).

1st and 2nd generation liquid biofuels have been subsidized largely on the premise that they are domestic substitutes for imported oil and that they can reduce greenhouse gas (GHG) emissions (Koplow, 2006). Numerous models indicate that biofuels are likely to play some role, and that some biofuel technologies can be effective at reducing carbon emissions in the long-term (e.g. Clarke, Jablonski, Moran, Anandarajah, & Taylor, 2009; Gül, Kypreos, & Barreto, 2007; Sarica & Tyner, 2013; Yeh, Farrell, Plevin, Sanstad, & Weyant, 2008). As we move forward into a world that requires increased decarbonization of fuel, it is imperative to better promote innovations in the biofuel landscape to decrease the carbon intensity of these fuels. The overall biofuel innovation narrative timeline is presented graphically below in Figure 3.1.

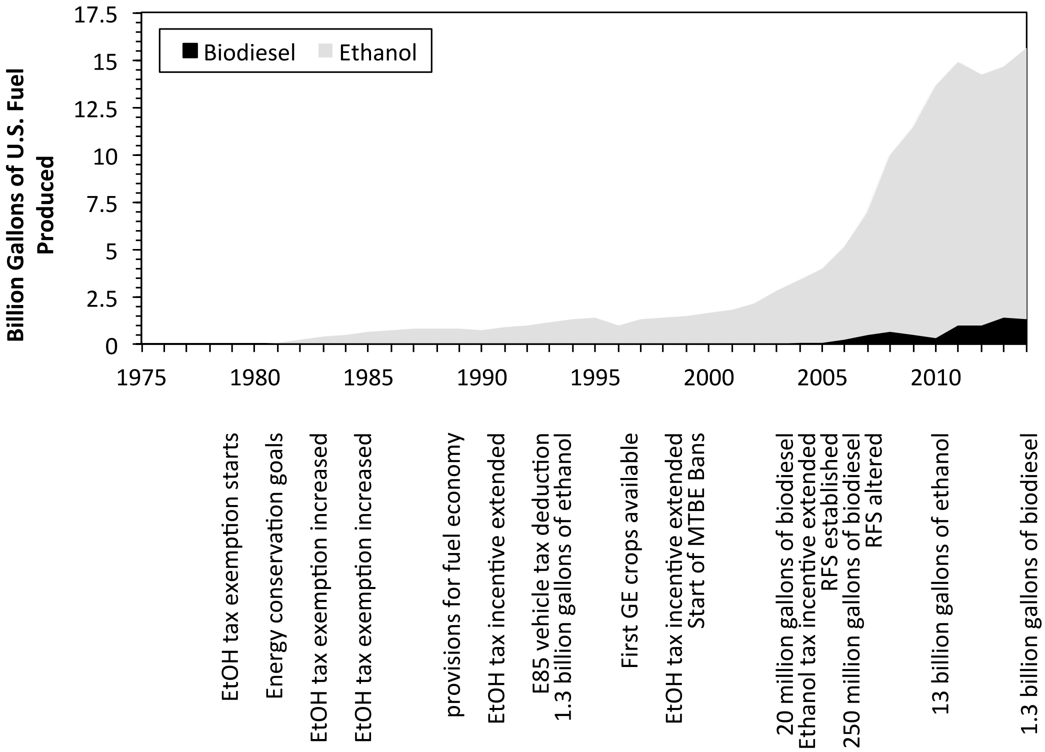

Figure 3.1. *Biofuel Innovation Timeline*

3.2 What Do the Patent Trends Tell Us?

To assess patent activity in relation to the U.S. Biofuel Narrative, I use three different patent classifications: (1) I use IPC codes and the Green Index for biofuels (GI), (2) I use the CPC classification scheme for biofuels (CPC), and (3) I use natural language processing approach to independently classify patents (NLP). I have made effort to compare patent classification across these three different methods, but not all classification systems can be used directly to track the same thing. The Green Index, for instance, does not facilitate direct comparison between 1st and 2nd generation technology. To compare Green Index patent activity, I use ethanol activity as a proxy for 1st generation biofuels. For the CPC classification scheme, I have combined Y02E50/16 (Cellulosic bio-ethanol) and Y02E50/18

(Bio-alcohols produced by other means than fermentation) to represent 2nd generation ethanol technology. Similarly, I assume Y02E50/14 (Bio-pyrolysis) is representative of 2nd generation biodiesel or renewable diesel. For classification and comparison purposes, I look at the filing year (application year) for accepted patents as opposed to the year that the patent was officially granted. Figure 3.2 shows the overall breakdown of results, where patent counts were normalized by the total number of patents that were filed in the United States in each year.

The order of magnitude across all classification methods is similar, with the exception being the modified, extended CPC classification (CPC*). CPC* takes the default CPC classifications for biofuels (Y02E50), and adds in an additional set of CPC classes to capture relevant biofuel patents that were not captured in Y02E50 – this can include things like feedstock and associated innovation.

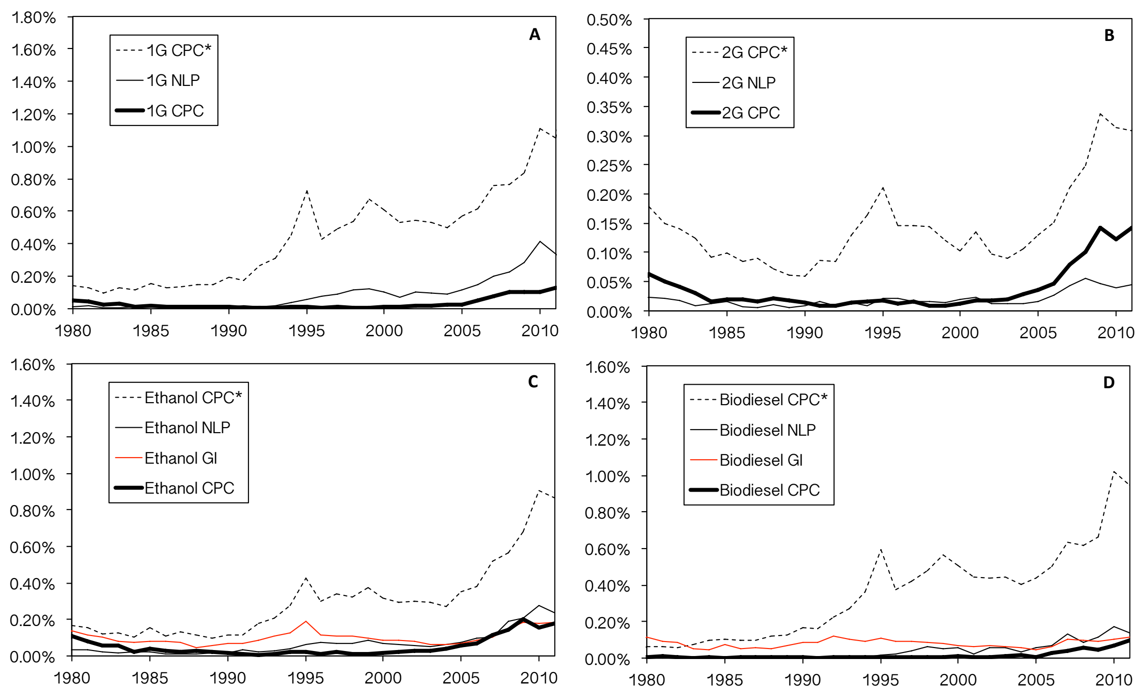

Figure 3.2. Biofuel patent trend comparison for Natural Language Processing (NLP) patent classification, Cooperative Patent Classification (CPC), Green Index (GI) patent classification, and a modified CPC-NLP patent classification (CPC*). Trends are shown for 1st generation biofuels (A), 2nd generation biofuels (B), ethanol (C), and biodiesel (D).

Furthermore, there appears to be one very consistent trend: starting in 2005, there was an increase in the amount of support for both 1st and 2nd generation biofuels. Prior to 2005, however, classification results seem to vary. I have curtailed patent data past 2011 as there is a 29-month lag time in the review of patents filed with the USPTO (U.S. Patent and Trademark Office, 2013). Given that this dataset was collected from 2013 through early 2015, it is likely that the number of patents recorded from 2012 onward is lower than the true number of filed patents, and therefore interpretation of trends or results past 2011 is unreliable. This long lag-time associated with the patent review process is one of the limitations in using patent data to assess innovation.

3.3.1 First-generation Biofuels

The NLP methodology for analysis of 1st generation biofuel-patenting activity (figure 3.1A) shows two distinct periods of patent filing (as sorted by application year). The first period from 1995 through 2000 was one of rapid expansion in biofuel patenting, followed by stagnation of activity, and then a second surge from 2005 through 2010. The existing biofuel narrative suggests that the first period of biofuel innovation was linked to innovation in genetically engineered food crops used in 1st generation technology processes, and that the second surge in innovation was driven by policies mandating and incentivizing the production and use of biofuels.

Although there was surge in biofuel patents from 1995 through 2000 in the NLP classification data, this result does not appear for the CPC or GI classification methods. Instead, GI ethanol data show an increase in patents leading up to 1995, followed by a gradual decline until 2005, while CPC data show a decline in patent activity starting in 1980, only picking up after 2005 – a contradictory result to the 1980s and onward narrative of a growing ethanol industry. Figure 3.3 shows overall 1st and 2nd generation trends for ethanol and biodiesel for the CPC and NLP classifications. When I modified the CPC classification scheme (CPC*) by including additional CPC classes, as selected based on the NLP-CPC concordance table, the data show a surge in patenting activity from 1995 through 2000

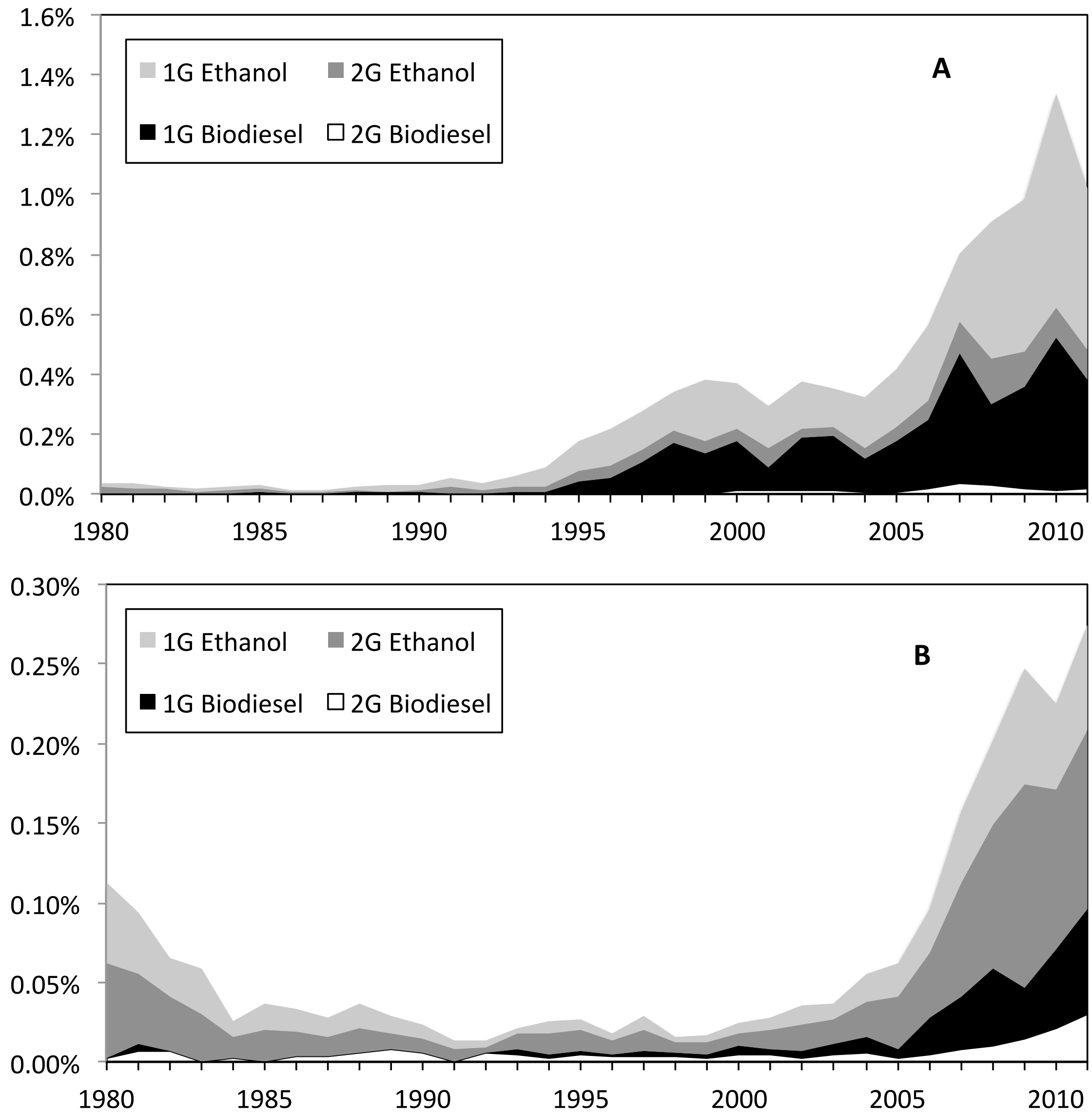

Figure 3.3. Biofuel patent trend comparison for (A) Natural Language Processing (NLP) patent classification and (B) Cooperative Patent Classification (CPC)

As GE crops proliferated in the 1990s, crop-related innovation spread across a variety of agricultural firms (Fernandez-Cornejo, 2009). New supply chains, networks, market conditions, and coalitions evolved that facilitated the later emergence of biofuel technology firms and the use of corn for ethanol (Fausti, 2015). Of the patent classification methodologies used, only the NLP-based approach corroborates this story, suggesting that support for biofuel in the 90s was aided by the agriculture sector and GE crops.

I utilize a general linear modeling approach to better assess the patent classification datasets to determine if 1st generation biofuel patent activity was linked to GE crop spillovers. Patent counts are sometimes zero, patent counts cannot be negative, and there is an overall exponential trend toward later years in the patent count data. I further checked the data for overdispersion, which turned out to be significant. Because of these properties to the data, I utilize a quasipoisson (QP) and a negative binomial (NB) model. Both of these models are well suited for dealing with count data, and accounting for overdispersion issues (Ver Hoef & Boveng, 2007). The following reduced-form model is used:

$$\ln\left(C_{1G,t}\right) = \beta_1 C_{All,t} + \beta_2 C_{Ag,t} + \beta_3 C_{BTech,t} + \beta_4 P_{corn,t-1} + \beta_5 P_{oil,t-1} + \beta_6 \{GE\}_t + \beta_7 \{RFS\}_t + a + \epsilon_t$$

Where C_{1G} is the 1st generation biofuel patent count, C_{All} is the count of patents in all sectors, C_{Ag} is the count of agriculture patents, C_{Btech} is the count of biotechnology patents, P_{corn} is the price per bushel of corn adjusted for inflation (National Agricultural Statistical Service, 2015), P_{oil} is the price per barrel of crude oil adjusted for inflation (McMahon, 2015), GE is a dummy value of 0 or 1 for whether or not GE crops are commercially deployable, RFS is a dummy value of 0 or 1 for the Renewable Fuel Standard being in place, a is the intercept and ε is the residual. I also added a 1-year time lag to oil and corn prices to allow time for any price effects to change patenting behavior. I used the Akaike Information Criterion (AIC) to compare models, and to assess the corn and oil lag structure (Sakamoto, Ishiguro, & Kitagawa, 1986). I looked at no lag, 1-year lag, 2-year lags, and 3-year lags for these variables. The AIC values indicated that a 1-year lag improved overall model fit compared to the other lag structures.

Results from both the QP and NB models are presented as there is no general answer for which model is the best to use (Ver Hoef & Boveng, 2007). The regression results in Table 3.1 indicate that RFS adoption and agriculture patents are significant variables for all models and patent data sets. The lagged price of corn also appears to be a significant predictor for 1st generation patent counts in the NLP dataset. Both the NLP and CPC datasets suggest that spillover effects from GE crop proliferation may explain some of the increase in 1st generation biofuel patenting activity.

Both the CPC and the GI patent classifications show biotech patents as a significant predictor of 1st generation patenting activity. The direction of the effect, however, is different for each classification set. This discrepancy highlights the importance in choosing correct patent classification approaches when it comes to interpreting results correctly. All three classification sets show different, confounding results for whether or not biotech patenting activity can explain 1st generation biofuel patents. Given that two classification models show different directions of effect, it seems likely that biotech-patenting activity is not a significant driver of 1st generation biofuel patents.

Only the NLP patent classification results remain consistent across both statistical models (negative binomnial and quasipoisson). The CPC results show a negative effect of corn prices on patent counts, as well as a negative effect of biotech patenting activity on biofuel patent counts, results at stark contrast with the GI patent classification and the NLP approach.

Table 3.1

Marginal effects from GLM regressions of 1st generation biofuel technology for three different patent classification methods

	NLP		CPC		GI	
	QP	NB	QP	NB	QP	NB
Other Sector Patents	-8.7E-05	3.5E-04	2.1E-04	2.8E-05	-1.1E-04	4.1E-06
Ag Patents	0.062 ***	0.04 *	0.014 .	0.019 *	0.068 *	0.049 .
Biotech Patents	2.1E-04	2.0E-03	-0.003 **	-0.0028 **	0.0076 *	0.0088 **
Corn Price	15.3 *	17.4 **	-3.0E+00	-5.2	29.8 **	24.5 *
Oil Price	4.1E-02	-2.8E-01	0.31 *	0.40 ***	4.1E-01	3.1E-01
GE Crops	81.532 **	42 *	1.3E+01	22.5 *	-4.2E+01	-3.3E+01
RFS	93.8 ***	115 ***	29.8 **	31.1 **	123 **	126 ***

Note. * = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$.

In 2005, a new wave of biofuel innovation began. I included an RFS dummy variable in the regression analysis to ascertain whether or not implementation of the RFS explains shifts in biofuel patent trends. Given the narrative of the dutch biofuel innovation system, in which 1st generation technologies originate from the agriculture sector (Suurs & Hekkert, 2009), I would also expect agriculture patenting activity to be significant in explaining patent trends for 1st generation technologies in the US. As shown in Table 3.1, implementation of the RFS appears to be extremely significant across all patent classification methods and both statistical models for 1st generation biofuel patent counts. Agriculture patents are also positive and significant for all patent classification methods, with increased statistical significance for the NLP classification data. This suggests that trends associated with agriculture patenting activity are similar to the trends associated with biofuel patenting activity.

Based on the biofuel technology narrative shown previously, it appears that patent activity, at least for 1st generation biofuel patents, aligns with many of the sociotechnical shifts occurring within the industry. The patent data support the narrative that spillovers from the agriculture industry affected biofuels, and that the RFS has been a significant driver of biofuel growth.

Of the patent classification methods chosen, the NLP methodology appears to provide results that are consistent, and that better align with the literature-based biofuel innovation narrative compared to the other two patent classification sets. The NLP-classified patents, for instance, specifically show an increase in patenting activity during the 90s, a period marked by shifts in agriculture and the rise in GE crop usage. The data also show a gradual build-up of patenting activity starting after 1980. In contrast, the CPC-selected biofuel patents are considerable starting in 1980, and go through a steady decline before surging again after enactment of the RFS (Figure 3.3). The activity portrayed by CPC patents is contradictory to the biofuel narrative, in which limited production of or interest in ethanol existed prior to enactment of ethanol policy. Additionally, the NLP approach is the only approach that aligns with the biodiesel innovation narrative: limited interest in biodiesel until 1992, where interest gradually increased and commercialization began later in the decade.

Using the NLP classification dataset, I also assess patenting activity for firms. The number of firms filing for 1st generation biofuel patents increased from less than 10 at the start of the 1990s to over 60 by the end of the decade (Figure 3.4). This is consistent with the biofuel policy at the time, where additional support for developing ethanol production facilities came from the small ethanol producer tax credit. This credit was first passed in the Omnibus Budget Reconciliation Act of 1990, and gave certain producers a 10¢/gallon credit on their first 15 million gallons produced each year. Plants with a nameplate capacity in excess of 30 million gallons a year were not eligible (Koplow, 2006). Correspondingly, Boone and Ozcan (2013) show a substantial increase in the number of small ethanol cooperatives that emerged after 1991, while the larger corporate-owned facilities stagnated until the start of 2005. This narrative continues to align with the patent data collected using the NLP classification methodology.

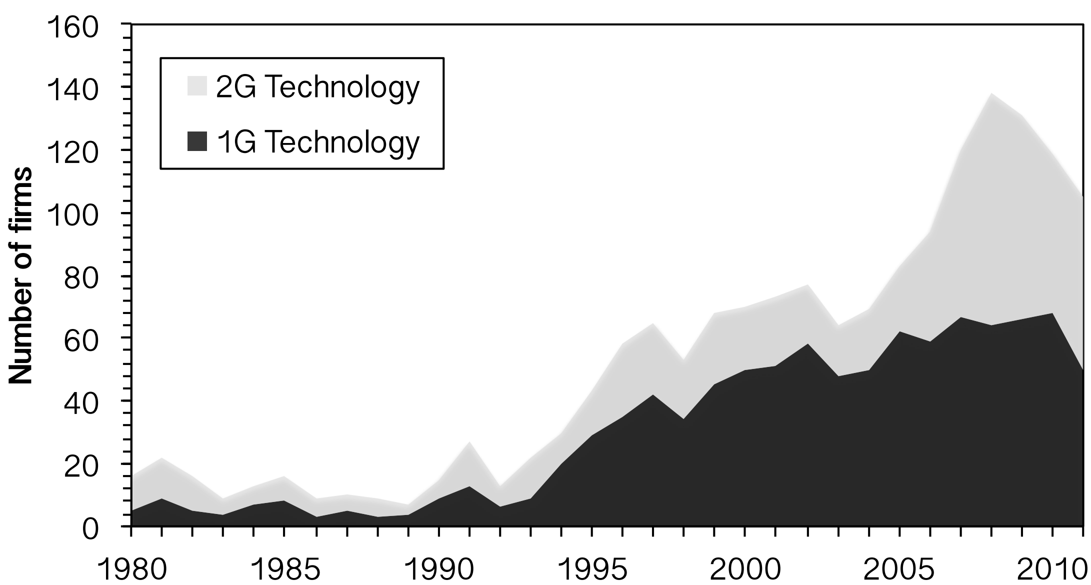

Figure 3.4. Number of firms filing for biofuel patents in a given year (classified using NLP)

Looking at which firms filed for patents in the NLP patent dataset, it appears that the dominant and established biofuel technology firms that emerged out of the agriculture sector in the 1990s continued to actively pursue food-related, 1st generation biofuel patents. According to the NLP-classified patent data, from 1990 through 2011 the top 5 firms filing for 1st generation patents were all agribusinesses – Mertec, Monsanto, Pioneer Hi-Bred, Stine Seed Farm, and Syngenta. These firms hold 61% of all 1st generation biofuel patents, and have filed an average of 45 biofuel-related patents per year.

3.3.2 Second-generation Biofuels

Prior to 2005, there was limited patent activity associated with 2nd generation biofuels for the NLP dataset (Figure 3.2B). Similarly, tangible policy support for 2nd generation biofuels only emerged with the RFS, and there was no production of 2nd generation biofuel until 2012. After enactment of the RFS in 2005, there was moderate growth in 2nd generation biofuel patents. 2nd generation technologies have been supported on the premise of improved environmental benefits and elimination of the food-vs-fuel issues associated with 1st generation technologies (Tyner, 2008, 2012). Further enthusiasm for the use of 2nd generation biofuels has come from the biotech industry, with support for the use of biotechnology as a means to address the low-carbon energy challenge (Lynd et al., 2008; Sanderson, 2006; Schubert, 2006).

To better assess 2nd generation biofuel trends, I employ general linear modeling. I use both a negative binomial model as well as a quasipoisson model to test CPC data and NLP-classified patent data. I use the following reduced form model:

$$\ln(C_{2G,t}) = \beta_1 C_{All,t} + \beta_2 C_{Ag,t} + \beta_3 C_{BTech,t} + \beta_4 P_{corn,t-1} + \beta_5 P_{oil,t-1} + \beta_6 \{GE\}_t + \beta_7 \{RFS\}_t + a + \epsilon_t$$

where C_{2G} is the patent count for 2nd generation biofuels. As shown in Table 3.2, both the enactment of the RFS as well as oil pricing were significant predictors of patenting activity for NLP-classified patents as well as for CPC patents. Implementation of the RFS starting in 2005 appears to have had a considerably diminished effect on 2nd generation patents compared to that for 1st generation biofuel technologies across both classification methods.

Table 3.2

Marginal effects from GLM regressions of 2nd generation biofuel technology for three different patent classification methods

	NLP		CPC	
	QP	NB	QP	NB
Other Sector Patents	8.9E-05	6.7E-05	1.2E-04	1.2E-04
Ag Patents	3.4E-03	2.6E-03	0.034***	0.031 ***
Biotech Patents	0.0014 *	0.0016 **	-0.0030 **	-0.0024 **
Corn Price	1.9E+00	1.1E+00	2.9E+00	1.8E+00
Oil Price	0.25 *	0.24 **	0.60 ***	0.64 ***
GE Crops	1.6E+00	3.6E+00	-6.3E+00	-6.1E+00
RFS	24.4 **	26.3 ***	27.5 **	28.1 **

Note. * = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$.

Disagreement between the two classification approaches, however, occurs when looking at the role of agriculture patents and biotech patents. For NLP-based classification, it appears that biotech patents are significant, positive predictors of 2G biofuel patenting activity whereas biotech patents are a significant, negative predictor for CPC patent activity. The NLP classification results support the idea that 1st and 2nd generation technologies have very different sectors of origin, whereas the CPC results indicate that agriculture patenting activity is still a dominant predictor of 2G patent trends. This once again highlights the fact that different patent classification schemes or groupings can result in radically different interpretations of events.

2nd generation biofuels have been promoted on the premise that they do not compete directly with food crops, and that they may offer further environmental benefits. CPC trends suggest that patenting activity in the agriculture sector is predictive of patenting activity for 2nd generation biofuels. Given this desire to decouple 2nd generation biofuels from traditional agriculture, the significance of agricultural patents as a positive predictor for 2nd generation CPC patent counts is concerning. The CPC data also indicates that more 2nd generation biofuel patents were filed compared to 1st generation patents during the 90s. Given the limited demand for biofuels in the 90s, and that 1st generation ethanol production stagnated prior to expanding during a period of MTBE bans, interpretability of CPC patent trends are suspect. NLP patent classification results suggest, however, that 2nd generation biofuels are significantly decoupled from trends in the agriculture sector. The patent data derived from the NLP approach better corresponds to the 2nd generation biofuel technology narrative obtained from the literature, where interest in 2nd generation biofuels was limited prior to the adoption of the RFS in 2005, and then its subsequent modification in 2007 resulting in direct support for advanced biofuel.

I again use the NLP patent dataset to assess patent trends for 2nd generation biofuel firms. Analysis of this dataset shows that patents came from firms more closely linked to the biotechnology sector (e.g, Novozymes, as opposed to Monsanto). From 1990 through 2011, the top 5 firms that filed for 2nd generation patents were Danisco, Genencor, Iogen, the Midwest Research Institute, and Novozymes. These firms filed an average of 2.4 biofuel-related patents per year.

The increase in 2nd generation biofuel patenting activity by biotech-related firms may be the same spillover phenomenon observed with agricultural firms and biofuel innovation in the 1990s: the opportunistic transfer of technology to a new sector. Following enactment of the RFS in 2005 and modification in 2007, firms originating from the biotechnology sector decided to experiment with biofuel technologies (Schubert, 2006). The number of firms filing for 2nd generation biofuel patents increased dramatically after enactment of the RFS. Conversely, the number of firms filing for 1st generation patents showed minimal growth (Figure 3.4).

While many firms filed patents for 2nd generation technology, the number of patents per company was much lower for 2nd generation technologies than for 1st generation technologies from 2005 onward. 76% of 2nd generation firms received only one patent in a given year. In contrast, only 3 to 5 firms were responsible for more than half of all 1st generation patents (almost 2,800 patents from 2005 through 2011).

3.3.3 Assessing patent quality

As discussed, patent filing activity varies for biofuels overtime, and NLP-derived patent trends seem to correspond to the biofuel innovation narrative obtained from the literature. However, without further consideration of patent quality, it is uncertain how the overall innovation intensity changes overtime. Traditionally, patent citation counts have been used to measure the importance of any given patent (Hall, Jaffe, & Trajtenberg, 2005; Lanjouw & Schankerman, 2004; Popp, 2001, 2005). The logic in using patent citations to assess patent value is that patents that are cited more often are likely to be patents of higher value. If, in a given year, there are fewer patents filed but there are a lot of citations in future years to those patents, then that may represent a high level of innovation for that technology. In turn, if many patents are filed but there are fewer citations, this may indicate that those patents are less innovative.

To assess patent quality, I have collected patent citation data for each NLP-classified patent in the database. As patents that have been around for a longer period of time have the possibility to gather more citations, I choose to look only at the number of citations linking to any given patent in a 2-year period from the year that the patent was granted. In other words, if a patent was filed in 1980 but wasn't granted until 1983, I look at the number of citations for that patent from patents that were filed in 1983, 1984, and 1985. Using this approach, I determine the quality of biofuel patents through 2009 (capturing citations from 2011 patents). The 2-year window for citation evaluation was chosen so that my patent dataset could be reliable for several years after 2007, the year in which the RFS was modified.

In addition to biofuel patent quality, I have also gathered patent citation data for the non-biofuel patents in the database (those categorized as "NA"). I use this citation data as an approximation for the overall trend in patent quality for all USPTO patents. To calculate average patent quality in each year, I took the sum of patent citations for each technology class associated with patents from a given filing year, then divide the total number of citations by the total number of patents applied for in that year. Figure 3.5 shows citation data for 1st and 2nd generation patents, as well as citation data for "NA" patents.

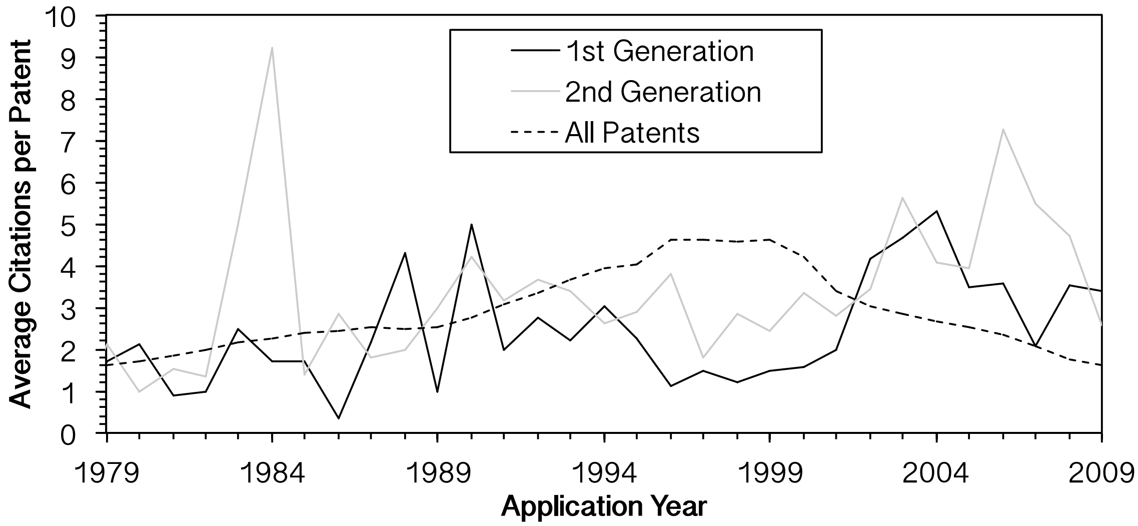

Figure 3.5. Patent citations per patent for 1st generation biofuel patents, 2nd generation biofuel patents, and non-biofuel patents in the NLP-classified biofuel patent dataset.

I utilize linear regression to better establish whether or not there is a time trend associated with patent quality. The following model is used:

$$ln(Q_s + 1) = \beta_1 ln(Q_{NA}) + \beta_2 t + \alpha + \varepsilon_{t,s}$$

where Q_s is the average 2-year citations count per patent in a given technology group s (1st generation or 2nd generation), Q_{NA} is the average 2-year citation count per patent for "NA" patents (average patents), and t is a linear time trend. I utilize the natural log of average patent citations because the average number of citations is greater than or equal to 0. 1 is added to Q_s to make zeroes that would otherwise occur in the count data non-zero. Regression results indicate that the time trend for both 1st and 2nd generation patents is insignificant (Table 3.3).

Table 3.3

Regression results to determine if there is a time trend associated with 1st and 2nd generation biofuel patent quality

	1st Generation	2nd Generation
(Intercept)	0.99 ***	0.66 **
$\log(Q_{NA})$	-0.65	0.74 *
Time	0.027	-0.013

Note. * = $p \leq .05$, ** = $p \leq .01$, *** = $p \leq .001$.

Additionally, I conducted a two-tail student's t-test to establish if patent quality is substantially different between 1st generation and 2nd generation patents. I find that the average patent quality between 1st and 2nd generation biofuel technologies is not significantly different. These results indicate that overall patent counts for biofuel technologies are likely a good indicator of general innovation activity in the system, and that the quality of patents has been fairly consistent over time.

3.3.4 The role of the oil and gas industry

Given how closely trends in NLP-classified patents align with the biofuel technology narrative, I explore the use of this data in providing more insight into the biofuel innovation process. Missing from the biofuel technology narrative, up to this point, is the role of the oil and gas industry. Since 1980, biofuels have steadily gained market share as a transportation fuel; this has been especially true since enactment of the RFS. Biofuel policy has created market opportunity for both agriculture and biotech companies, but the response from the oil and gas industry has been less certain.

The oil and gas industry could invest in advanced biofuels either as competitors to agricultural and biotech companies, or as partners. To date, oil companies have announced a myriad of investments into biofuels, often in partnership with large research universities (Sims, Mabee, Saddler, & Taylor, 2010). These incumbent fuel firms are rooted in petroleum production, distribution, and marketing. How does patent data portray the oil and gas industry? NLP-classified patent data show that the oil industry has not been a key contributor to knowledge development and diffusion for biofuels in either wave of biofuel innovation.

The number of biofuel patents filed by the oil and gas industry, normalized by the total number of U.S. patents filed in each year, is presented in Figure 3.6. Data are derived from the five largest oil and gas companies traded on the New York Stock Exchange in 2014 (BP, Chevron, Conoco, ExxonMobil, and Shell). On average, these five companies file for roughly 250 new patents each year. They filed for an average of less than one biofuel patent per year until 2006. But even from 2006-2010, oil companies filed fewer than 5 biofuel-related patents per year, representing less than 2% of their total company patents.

It might be that other oil and gas companies are more innovative in the biofuel space. To better assess this, I looked at the patent profile of Valero, a large oil refining company that has invested in and procured numerous 1st generation biofuel facilities. Valero has not currently been assigned any biofuel-related patents. The U.S. Energy

Information Administration (2014) indicates that Valero is currently the largest U.S. oil refiner.

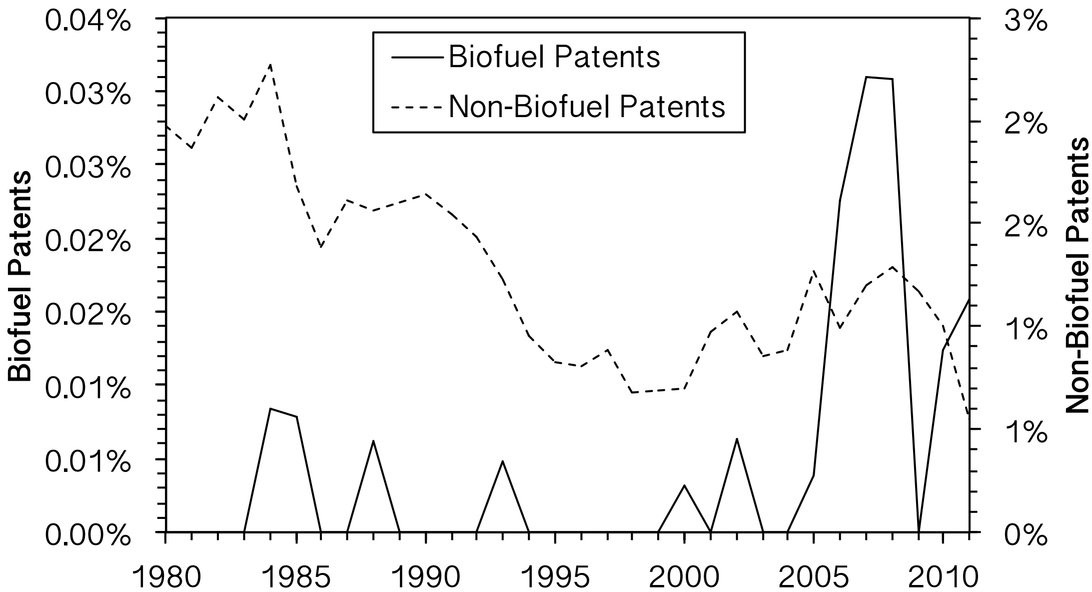

Figure 3.6. Biofuel patents and non-biofuel patents filed by the oil and gas industry as a percentage of all U.S. patents filed each year

Of the few biofuel patents filed by the five major oil and gas companies (N=42), all but one of these patents was for 2nd generation biofuels. These 2nd generation patents were mostly filed after adoption of the RFS. This pattern aligns with the existing 2nd generation biofuel narrative, in which companies pursued 2nd generation fuels after enactment of the RFS. The decrease in patenting activity by oil companies since 2009 may represent disinterest from the industry in biofuel advancement, an industry-wide preference for trade secrets rather than patents as a means to protect intellectual property associated with biofuels, and/or a preference for partnering with and investing in other smaller firms with the expectation of seeking licensing agreements with those small firms.

Patent activity has traditionally been viewed as an outcome of R&D expenditures, or as a means of approximating R&D expenditures (Griliches, 1998; Griliches, Pakes, & Hall, 1988; Kim & Marschke, 2004). The lack of a trend for biofuel patenting activity from the oil industry compared to other research endeavors by those same oil firms may be indicative of a lack of interest and commitment from the oil and gas industry toward biofuel technology. Given the limited biofuel patent activity from this sector, it is unlikely that biofuels will be strongly supported politically or financially by the oil and gas industry without substantially different policy approaches in the future.

3.3 Patents and Innovation

The patent analysis I conducted suggests that patent classification schemes and methods have a large impact on the interpretability of related technology innovation trends. Some patent classifications may act as a good indicator of technological shifts and overall innovation activity for technologies. Of the three patent classifications examined for biofuels, the NLP-classification methodology I utilized for classification seems to provide the most consistent empirical support for the biofuel innovation narrative obtained from literature.

While patent analysis is not without weaknesses, it appears that patent activity and trend, can in fact map on to the existing innovation narrative. This result tends to support the idea that patents can be used as a useful, readily available dataset for capturing some aspects of the technology innovation system. Despite these promising results, however, it is worth noting that the patent counts that I captured using typical patent classification methodologies did not always align with the biofuel innovation narrative. In some instances, trends were in direct contrast and provided contradictory results (a decline in patent activity with a growing ethanol industry). Due to this problem, I suggest that patent trends, derived from typical patent classifications, should be used with extreme caution. When using patents to construct an innovation narrative, these trends should not be used as the only source of information for evaluating the innovation narrative. If patents are used to assess technology innovation activity, additional effort should be made to adequately refine the results of patent searches to improve the overall dataset.

Even when using the NLP patent classification approach, the long lag-time associated with patent filing and the point at which a patent is granted is still a major drawback to using patent data. This makes it difficult for policy makers and modelers to understand what is occurring in the short-term, which makes it difficult to evaluate what can or should be done to improve innovation outcomes. To overcome this issue, I turn to a different data source: newspaper articles and trade journals.

Chapter 4: Text Classification and Biofuel Innovation

In this chapter I explore the use of textual media, specifically newspapers and trade journals, as a proxy for biofuel innovation. I employ natural language processing and machine-learning algorithms to classify newspaper articles. I have classified newspaper articles so that they directly map on to the Technology Innovation System (TIS) framework discussed in Chapter 1. Each article collected in the textual database represents information flow. Machine learning algorithms were employed to better determine the type of information flow associated with each article. Information was classified into the 7 TIS Innovation Functions (Table 1.2).

I find that trends associated with article counts closely correspond to the sociotechnical shifts described in the biofuel innovation narrative of Chapter 3. Additionally, I find that articles are a significant predictor of biofuel patenting activity for 1st generation, 2nd generation, ethanol, and biodiesel technologies.

Using a TIS-based approach to innovation assessment allows for increased disaggregation of innovation trends. To further assess the value of this approach, I examine the biofuel innovation case of California.

The TIS mappings and data indicate that the structure and support for innovation differs significantly in California compared to national trends. California's biofuel policies may explain many of these differences.

Qualitative assessment of the ethanol policy environment in California supports the idea that there is reasonable alignment between policy shifts and the innovation system trends as measured through textual media. I further employ statistical models to provide some support for the relationship between regional biofuel policy and shifts in the regional

TIS. The chapter finishes by looking at how support for different TIS functions across states might influence regional ethanol deployment and adoption.

4.1 Can Article Counts be used as a Proxy for Innovation?

As discussed previously, each TIS function represents a set of actions that can be taken to support innovation. Policy effects or actors and institutions operating in a technology innovation system can also directly support or block these functions. When sufficient support for TIS functions exists, structural components of the innovation system can fall into place.

For instance, loan guarantees or directed government grants could be provided to companies that decide to produce biofuel. These government interventions would encourage more entrepreneurial experimentation with biofuels, and would promote the additional entry of firms into the market. Ultimately, these structural changes could result in more knowledge creation and knowledge sharing about commercialization of the technology, which could lead to successful innovation. When adequate support for a function does not exist, the overall innovation system is weakened, which can have negative consequences for innovation. Determining what constitutes "sufficient" support, or what constitutes "weakness," in the system has largely been subjective, and can typically be evaluated only after innovation success or failure has been achieved.

The methodologies applied below may help to better identify weaknesses or strengths in a given innovation system as it is developing, which in turn can be used to aid policy makers that are trying to figure out how to better support or promote a specific technology or technology outcome.

To better understand biofuel innovation, hundreds of thousands of primary sources were captured and classified using the Stanford NLP Classifier to better map out and assess the technology innovation system for biofuels (this methodology is detailed in Chapter 2). A small subset of all texts was read in full, and each text was manually assigned to one or more of the 7 TIS functions. Each TIS function mapped on to an article was also classified as "supporting" the technology, or "blocking" the technology.

The Stanford NLP Classifier uses computational and statistical approaches to determine a set of article features (e.g. words, word pairs, article length, word prefixes and suffixes, etc.) that are statistically likely to be found in texts associated with specific classification. This allows for the classification of all articles into a set of supporting or blocking innovation functions. When aggregated together like this, it is possible to gain insight into what is or is not happening in a given TIS. The classified data used for analysis has further been constrained to only the United States due to the bias for English-language news sources for both collection and coding.

To establish regional news collections and specific geographies, I made use of the LexisNexis metadata that is available for each article that was downloaded from the LexisNexis article database. LexisNexis utilizes their own, proprietary algorithms for classifying articles. The metadata contained with each text document includes geographical information, as well as information associated with companies and other agents that are relevant to that article. The LexisNexis classification approaches are based on keyword filters and other article indexing techniques (LexisNexis, 2015). This metadata has allowed me to create an article subset corresponding to each U.S. State and Washington, DC.

I have chosen to characterize the entirety of the U.S. biofuel technology innovation system by aggregating each article subset associated with each state. Figure 4.1 shows a graphical representation of the trends and dynamics for the U.S. Biofuel TIS as exhibited by article counts.

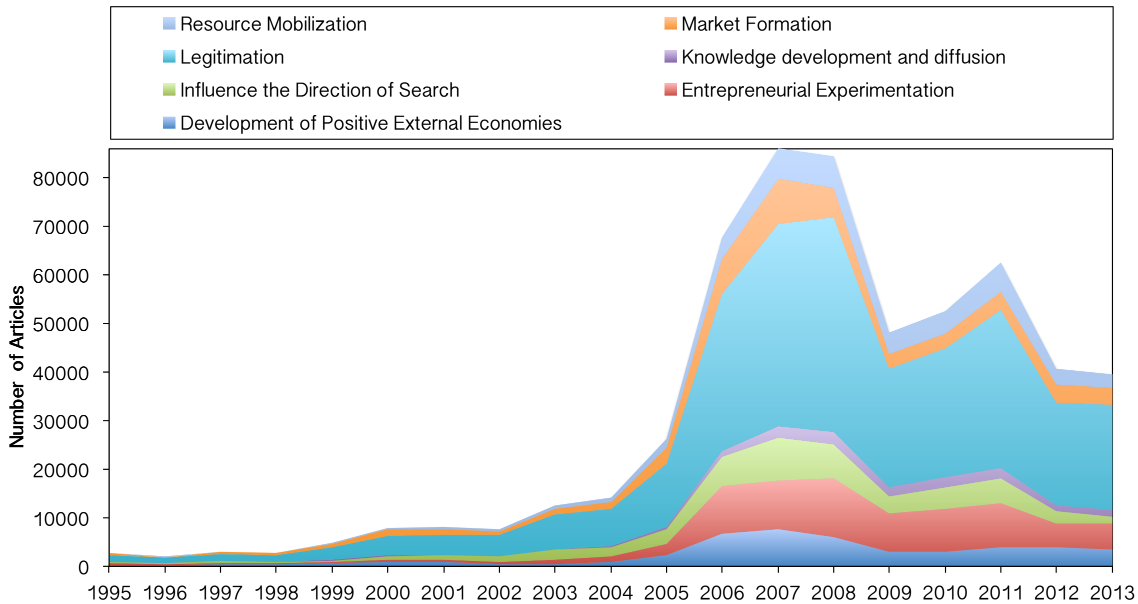

Figure 4.1. Overall biofuel technology innovation system trends for the United States

In Figure 4.1, there is a surge in the U.S. biofuel TIS from 2005 through 2008. This is a period characterized by the adoption and modification of the U.S. Renewable Fuel Standard. This sudden growth in the biofuel TIS is similar to what shows up in both the qualitative biofuel narrative obtained from the literature, a well as the patent data discussed in the previous chapter. The innovation system trends have an average coefficient of correlation of 0.82 with patent count data for ethanol, Biodiesel, 1G biofuels, and 2G biofuels. Additionally, Figure 4.1 shows an increase in articles starting in 1999 – this aligns with the narrative that the MTBE bans from 1999 onward may have provided some support for ethanol. Starting in 2008, and coinciding with the 2008 recession, the biofuel TIS decreases substantially. There is only partial resurgence of the TIS in 2011 before a greater decline in 2012 and 2013.

Compared to the patent trends discussed in Chapter 3, the economic recession is more pronounced in the article dataset. Although the number of patent applications exhibits a decrease starting in 2010 and 2011 for the NLP patent dataset, the number of firms patenting technologies shows a steep decline after 2008 (Figure 3.4) – this is similar to the trend exhibited by biofuel articles in Figure 4.1. The decrease in measured innovation activity in the article dataset due to the economic recession is expected, and directly aligns with the decrease in the number of firms filing for patents in the patent dataset.

To better assess the validity of article counts as a proxy for technology innovation trends, I compare TIS counts to the biofuel patent classification data of Chapter 3. I use general linear models for negative binomial as well as quasipoisson distributions to assess this:

$$\ln(P_t) = \beta_1 C_{All,t} + \beta_2 C_{Ag,t} + \beta_3 C_{BTech,t} + \beta_4 P_{corn,t-1} + \beta_5 P_{oil,t-1} + \beta_6 \{GE\}_t + \beta_7 \{RFS\}_t + \beta_8 A_t + a + \epsilon_t \quad \text{(M4.1 to M4.4)}$$

where P_t is the count of patents for years (t) 1995 to 2012, C_{All} is the count of all patents filed, C_{Ag} is the count of all agricultural patents filed, C_{BTech} is the count of all biotechnology patents filed, P_{corn} is the price of corn in dollars per bushel, P_{oil} is the price of brent crude in dollars per barrel, A_t is the article count for all TIS functions, GE is a dummy variable of 0 or 1 for the commercial deployment of GE crops, RFS is a dummy variable of 0 or 1 for whether or not the national renewable fuel standard is in place, a is the intercept and ε is the residual.

The model was assessed for each of the 4 technologies I have patent data and newspaper classifications for: (M4.1) 1st generation technology, (M4.2) 2nd generation technology, (M4.3) ethanol, and (M4.4) biodiesel. Patent count data was taken from the NLP patent classification dataset used in Chapter 3.

The GLM model results indicate that article counts for TIS functions are a significant, positive predictor of patenting activity across all biofuel technology patent subsets.

Of interest, when incorporating overall article counts to assess patent counts, the dummy variable associated with the implementation of the renewable fuel standard is no longer a significant predictor of biofuel patenting activity. This suggests that TIS article counts, as established through this methodology, may adequately capture policy implementation and other important sociotechnical shifts beyond what the dummy variables for the RFS is able to capture. For instance, different fuel volumes are required under the RFS each year, and various litigation activity alongside other market forces has likely had an effect on how well the RFS has been able to encourage technology innovation. This indicates that it is not simply a fuel mandate that encourages innovation, but that the specifics of the fuel mandate can have a large effect on innovation.

Table 4.1

Marginal Effects table for biofuel article classification models (M4.1 to M4.4) – marginal change in the number of patents per marginal increase of an independent variable

	1G Tech		2G Tech		Ethanol		Biodiesel	
	QP	NB	QP	NB	QP	NB	QP	NB
Other.Sector.Patents	-1.1e-3 ***	-1.6e-3 .	2.30E-04	2.3e-4 .	-1.2e-3	-1.8e-3 **	-4.9E-04	-4.2E-04
AgPatents	0.23 **	0.23 ***	-0.01	-0.01 .	0.15 **	.15 ***	0.11 *	0.11 ***
Biotech	-0.01	-.01 .	3.3e-3 *	3.3e-3 **	-2.4E-05	-4.6E-03	-3.8E-03	-3.5E-03
Corn_1	62 .	46 *	0.83	0.83	48 .	27 .	19	23
Oil_1	-0.21	6.00E-02	0.31	0.31 .	0.4	1	-2.8E-01	-0.41
GE_Rev	143	153 ***	-2	-2.0E+00	64	79 **	95	94 ***
RFS1	108	7.60E+01	2.3	2.3	82	56	27	22
Article Counts	2.4e-4 *	2.7e-4 ***	6.1e-4 ***	6.1e-4 ***	2.4e-4 .	2.5e-4 **	4.4E-04	4.9e-4 *

Note. .= p ≤ 0.1, * = p ≤ .05, ** = p ≤ .01, *** = p ≤ .001.

Given the relatively small time-series for the article dataset (1995 to 2013), I have conducted additional analysis using a pooled model, where I fit the model across all 4 biofuel technology cases (1G, 2G, Ethanol, and Biodiesel) to assess if article counts are a good predictor of patenting activity. I use a general linear model of the following reduced form for both quasipoisson and negative binomial families to assess this correlation:

$$\ln(P_{it}) = \lambda_t + \beta_1 A_{it} + \alpha + \varepsilon_{it} \qquad \text{(M4.5)}$$

where P is the number of patents associated with technology i for year t, λ is a fixed-effect for the year, A is the number of articles associated with technology i in year t, α is the intercept and ε is the residual. Akaike Information Criterion were also utilized to determine if there was a lag relationship between article counts and patents. I tested no lag, a 1-year lag, and a 2-year lag relationship. The no-lag relationship for article counts yielded the best model fit.

Regression results from the pooled model (not shown) indicated that article counts were once again highly significant for predicting biofuel technology patent trends. The fixed-effect for time (λ) was only significant in 1999, 2009, and 2010. Akaike Information Criterion suggests that further removing the fixed effect for time creates a better model specification.

Overall, regression results from the 5 different models employed suggest that article counts can, at the very least, be used as a proxy for biofuel patents. Because patents are often used as a proxy for innovation, article counts can also likely be used to assess technology innovation.

The overall number of articles related to biofuels, as shown in Figure 4.1, has diminished since the recession. This decrease in articles implies that support for biofuel innovation has faltered since peak levels in 2007 and 2008. Furthermore, there is substantial decline across all innovation system functions (classifications of article information) since the peak (Table 4.2). Additionally, there is no indication of an upward trend.

The greatest percent decline in innovation system functions for the national biofuel TIS occurs for the "Influence the Direction of Search" function. Statements of political support, creation of tax incentives, and reports of positive entrepreneurial outcomes characterize the "Influence the Direction of Search" function. All of these actions could motivate increased interest in biofuel technology. The substantial decline in the "Influence the Direction of Search" function may reflect overall uncertainty in pursuing biofuel innovation.

In the last few years, intensity has increased around the food-vs-fuel debate, and additional uncertainty has risen regarding the environmental benefits of both 1st and 2nd generation biofuels (G Cassman & Liska, 2007; Kendall & Yuan, 2013; Murphy & Kendall, 2015; Searchinger et al., 2008; Zilberman, Hochman, Rajagopal, Sexton, & Timilsina, 2012). Given these concerns, I would expect a decrease in the "Influence Direction of Search" innovation function since the peak of the biofuel technology innovation system.

Table 4.2

The year that articles associated with each innovation function peaked, and percent change in the number of articles from peak by 2013

Innovation Function	Peak Year	Change from Peak (2013)
Influence the Direction of Search	2007	-82%
Market Formation	2007	-61%
Entrepreneurial Experimentation	2008	-57%
Resource Mobilization	2008	-55%
Development of Positive External Economies	2007	-53%
Legitimation	2008	-51%
Knowledge development and diffusion	2008	-47%

To better assess what is occurring with 2nd generation biofuel technology innovation, I have aggregated the data to show innovation function counts for 2nd generation biofuel technologies. As seen in Figure 4.2, there is a large innovation system decline in 2012 and 2013 compared to the "biofuel peak." Unlike 1st generation technologies, however, 2nd generation technologies have experienced two peaks: one in 2007, and one in 2011, showing significant recovery after the economic recession in 2008.

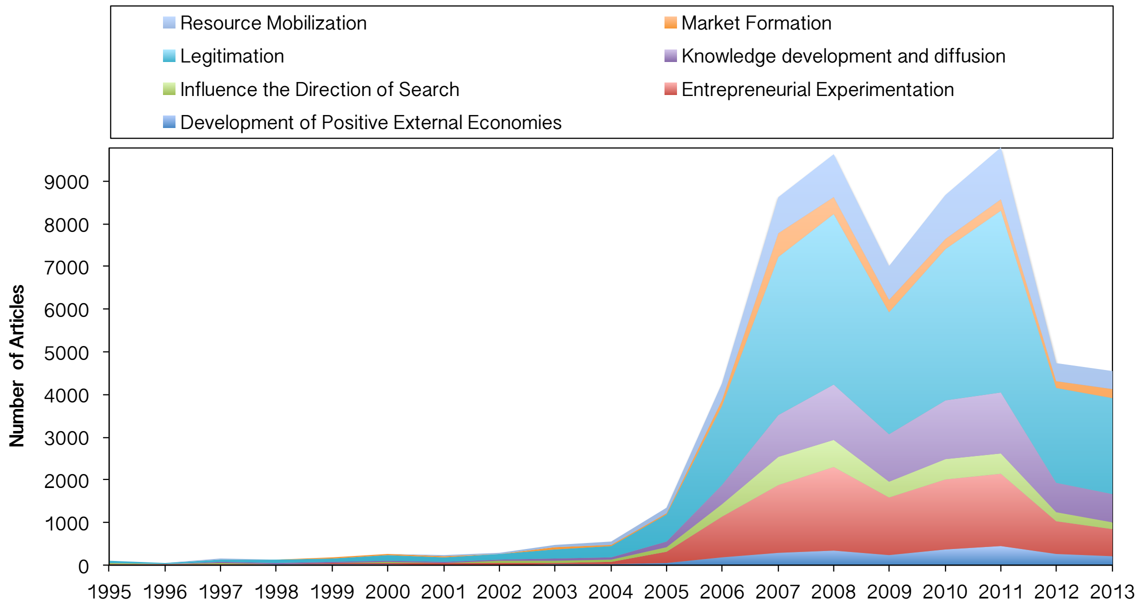

Figure 4.2. Overall 2nd generation biofuel technology innovation system trends for the United States

To better assess the "health" of the innovation system for 2nd generation biofuels compared to 1st generation technologies, I compare trends between the two technology innovation systems shown above. It appears that the percent decline has been greater for 2nd generation technologies since the peak of innovative activity than for 1st generation technologies. Entrepreneurial experimentation, resource mobilization, and knowledge development and diffusion for 2nd generation biofuel technologies have all declined considerably from their peak, and the magnitude of that decline is greater than for 1st generation technologies (Table 4.3). Given these large declines, improved innovation outcomes for 2nd generation biofuel technologies in the United States may be unlikely in the short-term.

Table 4.3

Innovation function peak year and percent change from peak by 2013

Function	1st Generation Peak Year	1st Generation Change from Peak	2nd Generation Peak Year	2nd Generation Change from Peak	Difference
Influence the Direction of Search	2007	-79%	2007	-75%	-4%
Entrepreneurial Experimentation	2008	-42%	2008	-67%	25%
Resource Mobilization	2011	-39%	2011	-65%	26%
Market Formation	2007	-60%	2007	-63%	3%
Knowledge development and diffusion	2007	-31%	2011	-55%	25%
Development of Positive External Economies	2007	-46%	2011	-52%	6%
Legitimation	2008	-40%	2011	-47%	6%

Although the results in table 4.3 indicate that the biofuel TIS at the national level has been in decline, it is not clear what is occurring at the state level, or what states have been doing to promote biofuel innovation independently. To better assess state-level innovation, I examine the case of ethanol in California.

4.2 The Case of Ethanol in California

California has a long history of environmental action; since the 1990s a number of new energy and environmental policies have been implemented in California, and legislative action has been taken to reduce GHG emissions (Franco, Cayan, Luers, Hanemann, & Croes, 2007; Schmidt, 2007). While California has done a lot to support environmental action, it was only in 1999 that biofuel, specifically ethanol, gained traction in the state with action to ban MTBE (Brekke, 2010). Over the next decade, ethanol utilization and production in California grew (U.S. Energy Information Administration, 2015c).

Starting in 2006, California implemented strategies to address climate change. The groundbreaking law AB 32, signed by Governor Arnold Schwarzenegger, called for a reduction of all GHG emissions by 25%, and California's Low Carbon Fuel Standard (LCFS) was initiated by Executive Order S-1-07 in early 2007 and adopted by the California Air Resources Board (CARB) as an AB 32 early action regulation in April 2009 (California Air Resources Board, 2011; Hanemann, 2007).

California's LCFS aims to reduce greenhouse gas emissions in the transportation sector by reducing the average fuel carbon intensity of transportation fuels sold in the state. Fuel carbon intensity is defined as the amount of GHG released through a fuel's lifecycle including extraction, conversion, transport and delivery, and consumption per unit of energy delivered. The standard is tightened over time to encourage and allow for cleaner fuel technologies to become more cost effective through innovation, economies of scale, and experience (Witcover, Kessler, Eggert, & Yeh, 2015; Yeh & Sperling, 2010). Policy like the LCFS is likely to impact a number of TIS functions. LCFS policy creates a protected market for low-carbon fuels (Market Formation). Additionally, credit trading mechanisms for the LCFS can be used to Mobilize Resources to support low-carbon technologies. Given that the LCFS is a long-term policy, with a transparent schedule for emission reductions, this increases the perceived legitimacy of the technology, and also aids in influencing the direction of search toward low-carbon fuel options.

Status reviews of the California LCFS show that compliance has relied on significant use of biofuel, specifically ethanol throughout the early phases of the program, and that a gradual shift away from ethanol to more advanced, lower-carbon biofuels and other low-carbon fuel pathways is occurring (Yeh & Witcover, 2014a, 2014b; Yeh, Witcover, & Bushnell, 2015; Yeh, Witcover, & Kessler, 2013). To date, California is the only state that has actively implemented an LCFS and has begun regulation. Oregon has implemented their own clean fuel standard, similar to the LCFS, but has not yet begun regulation, and the state of Washington was developing their own policy similar to California's LCFS before legislative action derailed implementation (64th Washington Legislature, 2015).

Due to the ability to differentiate newspaper event analysis by region in the TIS data set I have collected, I can directly compare technology innovation systems between regions. This, in turn, allows for interpretation of how implemented policy, at the regional level, may contribute to overall technology innovation outcomes. Given California's interest in driving reductions in GHG emissions, especially though utilization of market-based programs like the Low Carbon Fuel Standard, I would expect to see support for low-carbon technologies, such as biofuels, at a level that is significantly different from the national level. To test the significance of the innovation system trends seen above, counts for each article associated with a given TIS innovation function are utilized in a general linear model of the quasipoisson type. This accounts for overdispersion of newspaper articles, and random zero counts that occur. The following model is used:

$$\ln\left(A_{it}\right) = \lambda_t + R_i + \left(\lambda R\right)_{it} + \alpha + \varepsilon_{it} \qquad \text{(M4.6)}$$

where A_{it} is the article count for a given region in a specific year, λ is a fixed-effect for time, α is an intercept, and R is a fixed-effect for region. The logarithm is a link function, which is used primarily to keep the estimates of the response “in range”, given that all counts must necessarily be non-negative. Interaction effects across regions over time are considered. Figure 4.3 shows the biofuel TIS in California, and the smaller, embedded column chart shows which TIS functions are significant for California compared to the national trend for that function in a given year.

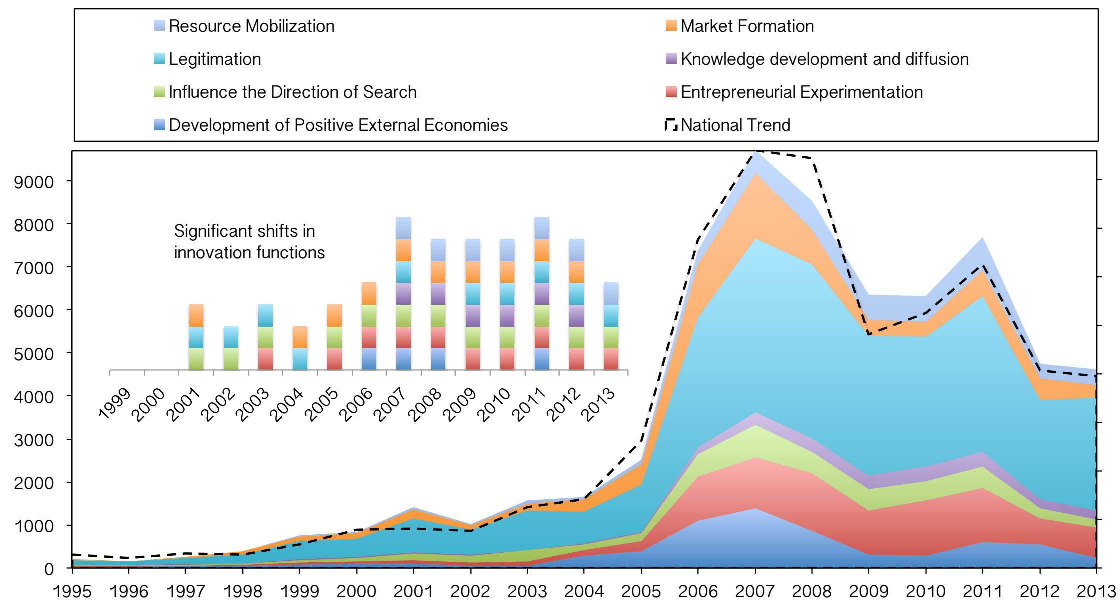

Figure 4.3. California technology innovation system trends for Biofuel compared to US trends

California has implemented a number of important, biofuel-related policies that can likely explain some of the significant shifts in the California TIS compared to what has occurred nationally. The first significant developments in California's biofuel TIS occur in the early 2000s, with additional support starting in 2006, a period when new GHG policy was being implemented. Results indicate that the system has undergone some decline in 2013. To better assess the role of policy in promoting the biofuel TIS in California, I look at one specific biofuel technology in California: ethanol.

Utilizing the above general linear model, I removed all technologies other than ethanol from the dataset, and again tested for significant deviations in California's TIS compared to other states. To assess the role of policy, I utilize the Department of Energy's Alternative Fuel Data Center (AFDC) policy database (U.S. Department of Energy, 2012).

This database provides information about the kinds of fuel policies that have been implemented over time, and what technologies each policy relates to (e.g. ethanol, biodiesel, etc.). In addition, the AFDC also provides classification terms for what kind of policy has been implemented (mandate, tax credit, etc.), and what user group is primarily affected by the policy (vehicle owners, station owners, etc.).

As seen in Figure 4.4, almost all significant shifts in the California ethanol TIS correspond to the implementation of new policy at the state level. The major exceptions are in 2010 and 2012. In 2010, no new policies were implemented. We see that the TIS does not change significantly since 2009 policy implementation, but still remains significant compared to the rest of the nation. Again, in 2012, there are few significant effects, whereas there was policy that was enacted. The enacted policy required that the California Department of Transportation develop and implement an alternative fuel vehicle parking incentive program. Effects from this program could be delayed, or AFV parking incentives did not significantly impact flex-fuel vehicle use or the ethanol TIS overall.

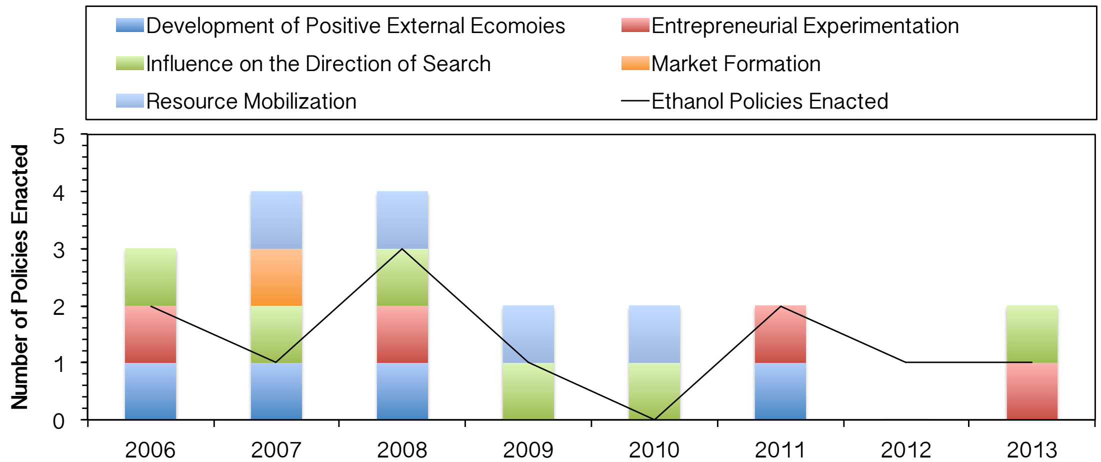

Figure 4.4. Significant shifts in California's ethanol TIS compared to the national ethanol TIS. The number of ethanol-related policies enacted in each year is also shown.

The results shown in Figure 4.4 allow for better comparison between my data-intensive approach to technology innovation analysis and policy expectations. This comparison fosters improved understanding of the policy environment, and facilitates assessment of the ongoing efficacy of certain policy implementations in terms of achieving specific goals.

Starting in 1998, California implemented an Alternative Fuel Tax, decreasing taxes associated with alternative fuels, including ethanol. Despite this tax credit being in place, no noticeable shifts in the California ethanol TIS occurred at this time compared to what was happening nationally. This tax was enacted during a time of on-going national discussion regarding the use of MTBE. With national shifts taking place at this time, a significant deviation from California's innovation system due to this policy would be surprising.

Given the ongoing national discourse associated with the use of MTBE, and that California implemented no additional policies that may have fostered the disuse of MTBE, the lack of significant innovation shifts is consistent with the California ethanol narrative.

The hiatus in both ethanol-supporting policy, as well as significant ethanol TIS developments in California compared to the rest of the nation continued until 2006, at which point California began enacting a number of GHG related policies, including AB 32 – with provisions for a clean fuel standard.

In 2006 and 2007, the California Energy Commission (CEC) developed an alternative fuel incentive program, allocating $25 million in incentives to promote the use of alternative fuels. In 2007 the CEC was further directed to adopt an annual investment plan. At the same time, provisions for a Low Carbon Fuel Standard began. In 2006 and 2007, considerable fuel policy was enacted in California, and we see a corresponding rise in significant innovation function support in Figure 4.4. Resource mobilization and entrepreneurial experimentation were supported by these California policies. Additional support came in the form of a protected market due to AB 32 provisions for a clean fuels program. Figure 4.4 results appear consistent with what was occuring in California in 2006 and 2007.

In 2008, California implemented policy associated with regulating vehicle retrofits, in addition to policy that encouraged state agencies to develop a plan for acquiring alternative fuel vehicles to reduce petroleum consumption; the intent was to reduce parking fees for alternative fuel vehicles, and to promote the use of alternative fuels when applicable to government fleets. The Low Emission Vehicle Standard implemented in the same year may have further influenced the direction of search toward ethanol, given the possibility for ethanol to reduce criteria pollutants. The shift in the TIS may be indicative of shifts in the national landscape to more directly promote ethanol. With adoption of the RFS, it is not surprising to see that California's market formation development for ethanol is insignificant compared to national trends.

With the recession impacting the biofuel innovation system across the U.S., California's innovation system appears to similarly falter after 2008. Ethanol production in California declined from 2.3 million barrels per year to 1.2 million barrels in 2009. Despite the recession, the California Department of Transportation was further tasked to develop a plan for achieving maximum feasible emission reductions. This requirement was imposed in 2009, and the final plan for implementation was slated for 2015. Even though plan development would take an extended period of time, long-term policies can provide some market certainty, and may further influence the direction of search toward fuels that can help to reduce vehicle emissions.

With economic recovery starting in 2010 and 2011, Propel Fuels, a private company, adopted their own station-wide policy aimed to encourage the use of ethanol in fleet vehicles (implemented in 2011). This policy provided small rebates to fleet vehicles fueling up with biofuel at their stations. This sort of privately funded policy aimed at increasing ethanol use is strongly indicative of entrepreneurial experimentation. During this time Ethanol production in California also surged, growing to 4.3 million barrels per year. Additional policy was also implemented by California at the same time to establish career training for students to prevent dropout rates, and to better promote interest in a clean economy – the development of a positive external economy.

2012 saw mandates to foster the development of parking infrastructure for alternative fuel vehicles for public parking areas operated by the California Department of General Services. Despite implementation of this policy, there were lawsuits associated with the LCFS and concern about indirect land-use change associated with ethanol. Compliance concerns with the Renewable Fuel Standard also emerged. The ethanol TIS in California does not appear to be better supported compared to the rest of the nation in 2012.

In 2013 California came forward in support of the National Renewable Fuel Standard, which was under jeopardy due to continued compliance issues. The California Legislature urged the U.S. Congress and the U.S. Environmental Protection Agency to take action to amend the Renewable Fuel Standard. This sort of action could have very well re-established support for ethanol and other biofuels in the state, promoting some renewed interest in California (influencing the direction of search).

While the above innovation narrative is not perfect, qualitative results indicate that there is alignment between the development and implementation of policy, and detectable shifts within the technology innovation system at the regional level. While more thorough analysis across a myriad of technology options and longer periods of time will be necessary to further validate this methodology, the initial analysis is encouraging, and provides a new means to assess policy efficacy for promoting innovation. Building on these results, I again turn to the regional level to establish which biofuel-related policies appear to have had an effect at encouraging innovation.

4.3 Assessing Biofuel Policy Efficacy

The Alternative Fuel Data Center (AFDC) tracks implemented state and national policies that affect alternative fuels and alternative fuel vehicles (AFVs). This database is maintained by the Department of Energy, and classifies policies into several different categories. The database includes information concerning when the policy was enacted, repealed, or amended, in addition to the primary stakeholders that the policy targets and brief policy descriptions. Although there are gaps in the dataset, it still provides an authoritative source for policies that were implemented at the state level. I make use of this policy database to assess which policy types correspond to significant shifts within a state's technology innovation system. Once again, I turn to the case of ethanol.

I look at the article counts associated with each innovation function separately. To assess the significance of policy implementations, I look at both the year that a policy was enacted, as well as the longevity effect of having the policy in place. I look at the interaction effect between the policy enactment date, the policy status, and the region over time. I use a mixed effect model of the poisson type to assess policy efficacy, allowing model error to be random. The following general linear mixed model is used:

$$\ln (A_{it}) = \beta_1 D_{it} * \beta_2 B_{it} * \beta_3 R_t + \lambda_t + 1|\varepsilon_{it} \qquad \text{(M4.7)}$$

where A is the article count for articles of a given TIS function, D is the dummy variable for whether or not a specific policy is in place for state i at time t, B is a dummy variable of 0 or 1 for whether the specific policy was enacted in year t for state i, R is a dummy variable for the region, and λ is a fixed-effect for time. The error term, ε, is a random effect. I only consider second order interaction effects between the policy dummies and the state. To assess policy implementation, I look at only one policy classification at a time. Table 4.4 shows which policies were significant when using this modeling approach.

Results from this model indicate that state-specific ethanol policies significantly predict technology innovation system trends. The policies that appear to have the greatest effect on the ethanol innovation system are Climate Change/Energy Initiatives (e.g. the LCFS in California) and Air Quality/Emission Regulation policy. State-level renewable fuel mandates also seem to encourage shifts in the overall innovation system.

Table 4.4

Policies that significantly affect innovation system article counts

Policies/Functions	F1	F2	F3	F4	F5	F6	F7
AFV Manufacturer Incentives			*				*
Fuel Use Incentives							
Aftermarket Conversion							
Air Quality/Emission Regulation	**				**	**	
Alternative Fuel Dealer Incentives	*						
Alternative Fuel Producer Incentives							
Alternative Fuel Purchaser Incentives			**			**	
Climate Change/Energy Initiatives	**				**	**	
Idling Regulation							**
Exemption from Restrictions							
Fleet Purchaser/Manger Incentives							

Fuel Economy Requirements							
Fuel Production/Quality Regulation		*					
Alternative Fuel Tax Rates							
Alternative Fuel Station Incentives							
Grants							
Loans and Leases							
Rebates		**			**		
Registration or Licensing Regulation	*	*			*	**	
Renewable Fuel Standard/Mandates	**				**	**	
Tax Incentives							
Vehicle Owner/Driver Incentive							

Note. F1: Resource Mobilization, F2: Market Formation, F3: Legitimation, F4: Knowledge Development and Diffusion, F5: Influence on the Direction of Search, F6: Entrepreneurial Experimentation, F7: Development of Positive External Economies
* Policy effects only occur after initial implementation
** Policy effects last during duration of implementation

To further assess policy's effect on the ethanol TIS, I use an additional modeling approach. Specifically, I use a negative binomial GLM model, taking into account the set of all (n) policy types as a predictor of article counts:

$$A_{it} = \sum_{j=1}^{n} \beta_j \left(Policy_j\right)_{it} + \lambda_i + \beta_{n+1} R_t \qquad \text{(M4.8)}$$

I include state (R) and year (λ) fixed effects for each TIS classification (i) and month (t). Using the Akaike Information Criterion (through the StepAIC algorithm implemented in R), I assessed all possible policy inclusions in the model to determine which policy dummy variables are most relevant for each innovation function. The marginal effects of policies that were significant for each innovation function are shown in Table 4.5

. These effects show the marginal increase in the number of articles associated with a specific innovation function that can be expected due to the implementation of policy. The percent magnitude of this effect is shown in parentheses. For example, if a region were to enact an AFV Manufacturing Incentive, the policy would effectively improve the Resource Mobilization function by 7%, but would have a larger effect on technology legitimacy (10% improvement).

Not all policies that were significant in Table 4.4 have corresponding values in Table 4.5, and vice versa. This is because M4.7 models what the effect of policy might be in isolation, while M4.8 takes into account the effects of several policies simultaneously, effectively estimating how policies, when they exist together, might influence the innovation system. For instance, from M4.7 climate change policy is significant in terms of Resource Mobilization, Influence in the Direction of Search, and Entrepreneurial Experimentation. However, in M4.8 results indicate that climate change policy only affects Entrepreneurial Experimentation. This is because climate change policy is likely to exist only alongside other policies that may also impact ethanol innovation. When considered with other policies, these other policies did not decrease model accuracy as much as climate change policy for predicting TIS outcomes. Other policy effects were more likely to exist across regions and therefore, when considered together, these policies better explain article counts than climate change policy alone.

Because it is very rare to find any single policy operating in isolation, there are interaction effects associated with every policy classification, and it is therefore difficult to parse out what effects can truly be attributed to only one type of policy.

Nonetheless, from this analysis, it becomes apparent that state-level policies have had some effect on the TIS as measured through article count data, and possess some explanatory power. Once again, there is indication that Renewable Fuel Standards/Mandates at the state level play an important role in the development of the ethanol TIS. The majority of policies appear to effect entrepreneurial experimentation, resource mobilization, or the pursuit of the specific technology (Influence in the Direction of Search). Both modeling approaches indicate that few policies directly legitimize technology or encourage basic R&D. This represents a fundamental gap in how states have formulated policies to promote technology innovation and adoption, and may indicate some weakness in the approaches that have been used at the state level to promote biofuel.

Table 4.5

Marginal Effect of Policy on Article Counts for Significant Variables

Policies/Functions	F1	F2	F3	F4	F5	F6	F7
AFV Manufacturer Incentives	2.5 (7%)		30.5 (10%)	1.1 (7%)	3.7 (6%)	5.0 (8%)	
Fuel Use Incentives	2.2 (6%)			0.7 (4%)			
Aftermarket Conversion						-1.9	3.3

						(-3%)	(9%)
Air Quality/Emission Regulation	-1.8 (-5%)						-2.7 (-7%)
Alternative Fuel Dealer Incentives		-3.2 (-6%)			-2.3 (-4%)		
Alternative Fuel Producer Incentives	1.3 (3%)					1.9 (3%)	
Alternative Fuel Purchaser Incentives		3.0 (6%)		0.7 (4%)			
Climate Change/Energy Initiatives						-2.8 (-4%)	
Idling Regulation		-3.2 (-6%)	-13.1 (-4%)	-0.8 (-5%)		-2.3 (-4%)	-2.6 (-7%)
Exemption from Restrictions					-2.8 (-5%)		
Fleet Purchaser/Manger Incentives	-1.7 (-4%)	-2.3 (-5%)		-0.9 (-6%)		-2.1 (-3%)	-2.2 (-6%)
Fuel Economy Requirements						5.1 (8%)	3.1 (8%)
Fuel Production/Quality Regulation				-0.6 (-4%)		-2.6 (-4%)	
Alternative Fuel Tax Rates							
Alternative Fuel Station Incentives	-1.3 (-3%)						
Grants		5.3 (11%)	-14.6 (-5%)				2.6 (7%)
Loans and Leases			-13.3 (-5%)				-2.8 (-7%)
Rebates							

Registration or Licensing Regulation			-11.2 (-4%)				
Renewable Fuel Standard/Mandates		8.6 (17%)					3.4 (9%)
Tax Incentives					3.4 (6%)	3.3 (5%)	
Vehicle Owner/Driver Incentive				-0.8 (-5%)			

Note. F1: Resource Mobilization, F2: Market Formation, F3: Legitimation, F4: Knowledge Development and Diffusion, F5: Influence on the Direction of Search, F6: Entrepreneurial Experimentation, F7: Development of Positive External Economies

4.4 Article Counts and National Ethanol Deployment

To close the loop and come full circle in terms of what article count data may be able to tell us, I look at the explanatory power of article counts on various deployment measures at the national level. Specifically, I look at ethanol production, ethanol consumption, flex-fuel vehicle (FFV) registrations, and E85 station counts to determine if article counts are a significant predictor of these factors.

I use several different modeling approaches to better assess deployment relationships. This is necessary due to the small size of the dataset. For many deployment terms, I only have annual time series data for each state from 1995 through 2010 (n=765). Specifically, there is concern with incorporating autoregressive factors, which are often exhibited for deployment (Cameron & Trivedi, 1986; Hurvich & Tsai, 1989). I try to reduce spurious results by aggregating individual TIS functions into an overall TIS system trend. In addition to using quasipoisson and negative binomial models, I also employ panel-based dynamic ordinary least squares (DOLS) to better capture the autoregressive nature of deployment. Judson and Owen (1999) show that dynamic panel models are likely to be effective when T > 30. Given my small dataset, I attempt to better compensate for consistency issues by employing robust standard errors in estimation. From this, I look at how innovation system trends affect deployment for 4 different deployment variables. The following reduced form model is used for DOLS:

$$\ln(D_{it}) = \beta_1 \ln\left(D_{i,t-1}\right) + \beta_2 \ln(Car\ Ownership_{it}) + \beta_3 \ln(Truck\ Ownership_{it}) + \beta_4 \ln\left(Innovation_{i,t-1}\right) + \beta_5 \ln(GDP_{it}) + \beta_6 \ln(Population_{it}) + \lambda_t + \alpha + \varepsilon_{it}$$

(M4.9)

where D is the deployment variable (flex-fuel vehicles, ethanol production, ethanol consumption, or E85 stations built) in region i for year t, Car Ownership is the number of new cars sold in that region that year, Truck Ownership is the number of light-duty passenger trucks sold in that region, Innovation is the aggregate article counts, GDP is the per capita GDP for a region in year t, Population is the population in that region, λ is a time fixed effect, α is the intercept and ε is the error term. The natural log of all terms was taken given that all values are greater than zero., 1 was added to the count of all logged variables to account for random 0s that occurred in the count data. Regression coefficients and significance from the four deployment models are given below in Table 4.6

Table 4.6

Regression results for ethanol deployment effects (Balanced Panel: n=51, T=14, N=714)

	Dependent variable:			
	FFV Deployment	Ethanol Production	Ethanol Consumption	E85 Stations
log(FFV.Lag + 1)	0.423***			
	(0.028)			
log(EthProd.Lag + 1)		0.900***		
		(0.025)		
log(EthCons.Lag + 1)			0.866***	
			(0.045)	
log(Station.Lag + 1)				0.534***
				(0.049)
log(Car.Ownership + 1)	0.581**	-0.615**	0.325	0.079
	(0.231)	(0.296)	(0.593)	(0.227)
log(Truck.Ownership + 1)	0.036	0.744*	-0.078	-0.185
	(0.116)	(0.387)	(0.669)	(0.202)
lag(log(Innovation + 1), 1)	0.042	0.141***	0.057	0.064**
	(0.041)	(0.053)	(0.112)	(0.032)
log(Economy)	0.615**	-0.665	0.671	-1.334***
	(0.267)	(0.634)	(1.633)	(0.500)
log(Population)	-0.717	1.344	-1.141	1.095
	(0.473)	(1.727)	(3.088)	(0.830)
year1998	3.135***	-0.186***	0.415	0.036
	(0.152)	(0.056)	(0.336)	(0.069)
year1999	2.975***	-0.256**	-0.008	0.047
	(0.157)	(0.113)	(0.257)	(0.079)
year2000	2.786***	-0.152	-0.017	0.166
	(0.179)	(0.189)	(0.351)	(0.110)
year2001	2.120***	-0.390*	0.046	0.174

	(0.210)	(0.201)	(0.477)	(0.134)
year2002	3.099***	-0.305	0.426	0.244*
	(0.194)	(0.229)	(0.506)	(0.147)
year2003	2.908***	-0.196	0.040	0.287
	(0.218)	(0.292)	(0.553)	(0.181)
year2004	2.382***	-0.509	0.252	0.542**
	(0.245)	(0.393)	(0.685)	(0.210)
year2005	2.337***	-0.438	2.152**	0.924***
	(0.237)	(0.411)	(0.836)	(0.256)
year2006	2.503***	-0.453	0.830	0.979***
	(0.244)	(0.424)	(0.867)	(0.296)
year2007	2.977***	-0.033	0.843	0.908***
	(0.280)	(0.400)	(0.970)	(0.321)
year2008	2.643***	0.145	1.122	1.058***
	(0.307)	(0.462)	(1.047)	(0.333)
year2009	2.797***	-0.173	0.865	0.701**
	(0.300)	(0.393)	(0.996)	(0.333)
year2010	3.213***	0.010	0.973	0.923**
	(0.338)	(0.455)	(0.988)	(0.360)

Note: *p<0.1; **p<0.05; ***p<0.01

From the DOLS regression results, we see that the state-level innovation system is a significant, positive predictor of E85 station rollout and ethanol production. Ethanol consumption does not show the innovation system as being significant. The lack of significance for the innovation system in this instance can likely be explained by the presence of national policies that dominate the effect of state incentives. The lack of significance of the innovation system in FFV deployment is unsurprising, given the positive significant effect of car sales. It seems likely that general car ownership and purchasing trends ultimately determined FFV deployment, a result inline with FFV purchasing behaviors, where there have been few compelling reasons for active adoption of FFVs (Collantes, 2010; Keefe, Griffin, & Graham, 2008).

I also used quasipoisson and negative binomial models to assess innovation system effects on deployment variables. These models take a similar form to M4.9, with the exception that state-based fixed effects (R) are included:

$$ln(D_{it}) = \beta_1 D_{i,t-1} + \beta_2(Car\ Ownership_{it}) + \beta_3(Truck\ Ownership_{it}) + \beta_4\left(Innovation_{i,t-1}\right) + \beta_5 GDP_{it} + \beta_6(Population_{it}) + \alpha + \lambda_t + R_i$$

For these models, the algorithm used to solve for the negative binomial coefficients was unable to converge for ethanol production or ethanol consumption. Testing for overdispersion across all models revealed that overdispersion was likely, indicating that a simple Poisson model would not be adequate. The marginal effects for the quasipoisson and negative binomial GLMs (where applicable) are given in Table 4.7 (not the beta coefficients).

Looking at Table 4.7, there is strong indication that car ownership was a significant predictor of FFV deployment, ethanol production, and ethanol consumption. The innovation system trends are shown to be significant and negative in the negative binomial model for FFV deployment, but not the quasipoisson model. For ethanol production and consumption, the quasipoisson model indicates that the innovation system is a significant, negative predictor of deployment. These results are largely contradictory to the results from M4.9.

Due to the large dispersion parameter for the data, the marginal effects for these models may be inaccurate. Additionally, the innovation system trends have a negative relationship with ethanol production, which is difficult to interpret. Given the contradictory results between the DOLS model and the QP/NB models, these regressions results should be treated with caution. Nonetheless, results across all models show that the innovation system has likely been significant in promoting E85 station rollout.

From this analysis, it is not clear if innovation system trends have significant predictive power for deployment outcomes. Additional lag effects may need to be considered, and alternative model formulations should be explored. For some deployment data, it may be possible to provide more resilient statistical analysis by taking into account monthly values. Ultimately, more data is necessary to establish the validity in the use of innovation system trends for predicting biofuel deployment outcomes. Data limitations aside, initial results suggest that there may be some relationship between article counts and deployment (such as for E85 Stations). To further investigate the use of these methodologies for predicting deployment outcomes, and for assessing policy, I turn to the case of an emerging technology: electric vehicles.

Table 4.7

Marginal effects and significant for deployment variables using quasipoisson (QP) and negative binomial (NB) general linear models (N=764)

	FFV Deployment		Ethanol Production	Ethanol Consumption	E85 Stations	
	QP	NB	QP	QP	QP	NB
Autoregressive	0.03 ***	0.036 ***	-1.7E-07	2.0E-01 ***	3.5E-02 ***	5.4E-02 ***
Car Ownership	0.015 ***	0.016 ***	4.3E-07 ***	6.7E-03 **	-3.4E-06	-5.3E-06 .
Truck Ownership	-0.006 ***	-0.007 ***	-1.3E-07	1.9E-03	8.3E-07	2.3E-06
Innovation_t-1	-0.005	-0.196 **	-5.0E-06 *	-2.5E-01 ***	1.7E-04 **	2.0E-04 **
GDP	-0.003 ***	-0.001	-3.9E-07 ***	-6.9E-03 ***	1.5E-06	1.1E-06
Population	-3.7E-06	-2.1E-04	1.1E-07 ***	1.5E-03 ***	1.0E-07	2.3E-07

Note. .= p ≤ 0.1, * = p ≤ .05, ** = p ≤ .01, *** = p ≤ .001.

Chapter 5: Emerging Technologies – The Case of EV Innovation

To better understand the innovation system for emerging technologies, I turn to electric vehicles. Given the still-developing nature of this technology, overall market dynamics remain uncertain, and a cohesive narrative for EV deployment and innovation remains lacking. To aid in developing the narrative for EV innovation, I analyze the EV market in three states: Washington, Colorado, and Georgia.

Meetings were orchestrated with 31 different individuals across these 3 different states. These individuals were from a variety of organizations and groups and were chosen based on their high profile status in the electric vehicle space within each respective state. Initial contact points were established through already-existing UC Davis state contacts that have participated in UC Davis's Zero Emission Market Acceleration Program (ZE MAP). The ZE MAP participants are believed to have good understanding of the network of actors and agents working on EVs in their state. Suggested individuals provided by these initial ZE MAP contacts were emailed to setup a 45-minute in-person meeting. Meeting participants, in turn, were also asked to provide a set of additional contacts that they found to be important in the EV space within the state. Most additional contacts had already been given by the ZE MAP point people, and were therefore already participants in the meetings.

Meetings were used to better gain insight into the EV technology innovation process. The information collected from these meetings has been used to assess commonalities and differences across these 3 different states. These discussions helped provide detail for the innovation narrative to better determine what factors promote or detract from EV adoption. EV Sales for each of the three states are shown in Figure 5.1.

There has been considerable, positive growth in the EV market for each of these states. While uptake trends are positive, the uptake path for each state is different.

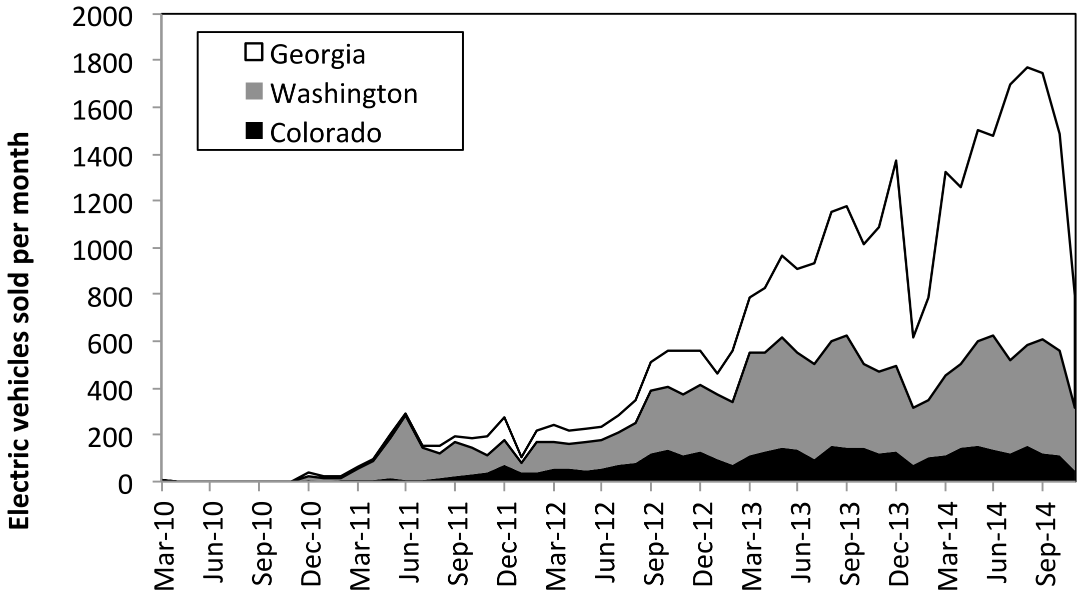

Figure 5.1. Electric Vehicle Sales in Washington, Colorado, and Georgia.

5.1 The Case of Washington

Washington is often included alongside California and Oregon during discussions of environmental policy and regulatory action. Washington has made many attempts to adopt policies and approaches similar to those in California to deal with greenhouse gas emissions. For instance, Washington recently had plans to adopt and implement a Clean Fuel Standard similar to the Low Carbon Fuel Standard in California (R. Brown et al., 2015). These implementation attempts, however, have been derailed (64th Washington Legislature, 2015). Furthermore, implementation of a Zero Emission Vehicle Mandate (ZEV Mandate) has also been discussed for Washington. The ZEV Mandate is a policy approach pioneered by California, and also adopted in the Northeastern States and Oregon, that works to promote the adoption of electric and fuel cell vehicles (Collantes & Sperling, 2008). Washington, unlike California and Oregon, however, has not adopted a ZEV mandate. This makes EV deployment in Washington uniquely different compared to either Oregon or California.

I utilized a statistical model (a negative binomial general linear model with a log linking function), in conjunction with monthly EV sales data for 24 different states from January 2010 through November 2014 to test for the significance of state-based fixed effects. The model was formulated as:

$$
\begin{aligned}
ln\{Vehicles\ Sold\}_{it} &= \beta_1\{Vehicles\ Sold\}_{i,t-1} + \beta_2\{Electricity\ Prices\}_{it} + \beta_3\{Rural\ VMT\}_{it} \\
&+ \beta_4\{Gas\ Price\}_{it} + \beta_5\{Temperature\}_{it} + \beta_t\{Year\}_t + \beta_r\{State\}_i \\
&+ \beta_{rt}\{State{:}Year\}_{it} + \varepsilon_{it} + \alpha
\end{aligned}
$$

(M5.1)

where ε is the error term and α is the intercept for each region (i) and each year (t). I included retail electricity prices, the per capita number of miles traveled on rural roads, and the average monthly temperature in each state as possible explanatory variables for EV sales (urban VMT was shown to be insignificant during model selection).

From modeling results, I find that the state-year fixed effect for Washington was not significant in determining the number of vehicles sold each month. The Washington fixed effect is slightly significant at only a 90% confidence level. That is to say: Washington has not significantly improved EV uptake since 2010, but Washington has potentially supported EVs better than other states for which I have deployment data.

Through numerous in-person discussions, I explored the critical factors driving this uptake. For Washington, participants stated that the key contributors for or detractors from EV uptake were: **price, limited range, and limited information**. Important to note is that during meetings in Washington, EV buyers were still described by meeting participants as "early adopters," that is, people of higher affluence that are more interested in experimenting with new technologies. This description is in stark contrast to how EV adopters were described in Georgia. Georgia stakeholders stated that EV adoption was not restricted to early adopters, but was instead motivated by "commuters."

In Washington, meeting participants often described EV sales as being driven by "word of mouth" and familiarity with seeing electric vehicles around Seattle, in addition to the compelling sales tax exemption that extended to **leased** electric vehicles. The early EV adopters were described as being technophiles, and individuals that wanted to play with a "new gadget," and therefore viewed a two-year or three-year leasing arrangement as an advantage.

Unlike most other states in the United States, Washington does not collect income tax – instead, public revenue is collected through use of a sales tax at the point of purchase. Given the lack of individual tax liability, Washington has decided to reduce the cost for purchasing electric vehicles by exempting EVs from the required vehicle sales tax. The sales tax exemption can amount to almost 10% of the price of the vehicle (Washington Department of Revenue, 2015). The sales tax exemption has recently been modified such that it only applies to EVs that cost $35,000 or less – thus directly excluding Tesla and BMW EV models from the scope of this incentive.

The sales tax exemption also applies to lease rates, albeit the "value" is lower than would otherwise be obtained through the outright purchase of the vehicle. Some meeting participants stated that consumers were unwilling to commit to fully purchasing the vehicle, and preferred the opportunity for a three-year lease. Leases allow consumers to obtain newer, better electric vehicle at the end of the lease. This may be important for EV deployment given the newness of the technology, and that improvements with the technology are happening quickly. For instance, the Nissan Leaf 2011 model had an estimated range of only 73 miles compared to the 75 mile range of the 2013 model, and 84 mile range of the 2016 model (Environmental Protection Agency, 2015).

In addition to the sales tax exemption, Washington has also been a leader in deploying EV Supply Equipment (EVSE), particularly charging infrastructure under the American Recovery and Reinvestment Act of 2009 (Powers, 2014). Additionally, the Washington Department of Transportation has put forward some funding toward EVSE.

Although there are numerous incentives for EVs in Washington, recent legislation has been enacted to tax electric vehicles for a road-use fee. This compensates for the exemption that EVs otherwise would receive from gasoline tax. The fee is $150 in the first year, and $100 each year afterwards. $50 from the first year goes toward supporting additional rollout of EVSE.

Most participants that commented on the road-use fees indicated that users did not like paying them, but thought that it was a decent plan for raising additional funds to support EVSE. Unlike in other states, meeting participants in Washington expressed concern about providing EVSE for multi-family homes and apartment complexes within the greater Seattle area. Other states had mostly discussed the rollout of EVSE to workplaces, and did not seem focused on multi-family homes at this phase of deployment.

As EVs have become more prevalent in Washington, Investor-Owned utilities in the region have started to consider rollout of their own charging infrastructure and network. Some meeting participants expressed concern over the ability for utilities to rate base infrastructure costs, and this uncertainty was attributed to the limited deployment of EVSE from utilities.

While the Washington sales tax exemption has applied to a myriad of EVs, sales of the Nissan Leaf have accounted for more than 72% of pure EVs sold in the Seattle area over the timeframe for which I have data. From meeting with individuals in the Seattle area, it became apparent that Nissan Leaf deployment in and around Seattle has been extensively supported by a "champion" dealer in the area. One dealership (more specifically, one specific salesperson at the dealership) has fostered and supported adoption of the Nissan Leaf far better than other dealerships. These "champion" dealers are well-versed in electric vehicles, being users of the technology themselves, that are readily able to answer questions and abate concerns that potential new-car buyers have when it comes to owning an electric vehicle. During meetings, one specific dealer in the Seattle area was regularly referenced as a major EV champion in Washington, being responsible for a majority of Leaf sales in the area.

5.2 The Case of Colorado

From meeting with a variety of important EV-actors in Colorado, a narrative took shape that revealed the perceived innovation system and network of important actors within the state.

Colorado has taken an aggressive approach to provide monetary incentives for electric vehicle adoption that far exceeds what many other states are doing. There is a fully refundable income tax credit in place that can provide as much as $6,000 in credit for the purchases of an electric vehicle. This amount is determined by the size of the battery in the EV, and requires a formula to calculate the true value of the incentive depending on the EV being bought. In addition to this strong consumer purchase incentive, EV outreach and education efforts were brought up in meetings far more often in Colorado than the other states I looked at (Washington and Georgia).

Meeting participants indicated that many EV stakeholders, ranging from owners and enthusiasts to dealers and government officials, regularly participated in ride-and-drive events. Ride-and-drives are events where the public can drive and ride in electric vehicles, gaining first-hand experience with the technology. In meetings, it was stated that these ride-and-drives have especially been targeted at workplace environments. Additional educational outreach was supported by Drive Electric Northern Colorado (Fort Collins and Loveland), and the Colorado Energy Office has supported events throughout the State.

In meetings, it was brought to my attention that a variety of new initiatives have started with intent to expand access to workplace charging infrastructure. Aggressive support has come from the Boulder Nissan Leaf dealership for charging infrastructure, where for every two new Leafs purchased by employees of a given company, a new charging station is installed. More targeted efforts by state and local governments to expand funding and grants for workplace charging infrastructure and fleets are also in place or planned.

Despite Colorado's refundable tax credit and electric vehicle supply equipment (EVSE) deployment efforts, EV sales have not been statistically significant in Colorado compared to elsewhere in the United States over the period of time for which I have data. Modeling results from M5.1, show that neither the state fixed effect for Colorado, nor the state-year fixed effect for Colorado was significant in determining the number of vehicles sold each month. That is to say, Colorado's EV sales have not been significantly different from what would be expected elsewhere in the country, and there have been no sudden annual changes from this expectation in the timeframe for which I have monthly data (2010 through 2014). Average monthly temperature was also found to be insignificant for determining EV uptake.

Participants indicated that they thought HOV benefits could further help promote EV adoption, and in meetings, participants expressed dismay that no high-occupancy vehicle (HOV) incentives existed for EVs in Colorado. Another factor often brought up in meetings that could possibly be inhibiting EV sales was "Colorado culture", where consumers demanded four-wheel drive (4WD) vehicles. Although, the 4WD or all-wheel-drive (AWD) demand for EV models may be a unique cultural phenomenon to the Colorado consumer market, the lack of model availability is likely not a "true" barrier, and instead only acts as a superficial deterrent that could be overcome with active dealer and consumer education. For instance, AWD and 4WD vehicles do not offer improved safety compared to front-wheel drive vehicles (Gårder, 2014). In one meeting, a stakeholder specifically mentioned that the 4WD or AWD barrier was not a true detractor from EV sales, indicating that modern traction control systems, anti-lock breaking, and tire technologies (snow tires) are capable of providing performance and safety conditions on par with or in excess of 4WD and AWD vehicles during winter months.

Instead, what I find from communicating with representatives in Colorado is that, despite educational outreach efforts, information is not properly being disseminated to the correct sources, and dealers have not received proper training or knowledge about EV incentives or about EV technology. Some meeting participants, including dealers, indicated that dealers are cautious when it comes to promoting EVs – possibly overly so. In meetings, participants suggested that the individuals buying EVs came into dealerships often knowing more than the sales representatives at the dealership about EV technology, tax incentives, and charging infrastructure.

Some dealers that I met with expressed frustration with the complexity of the Colorado income tax credit, and were not always aware of informational resources that had been created by government organizations to show the eligible tax credit amount for different vehicle models in an easy-to-access format. Many meeting participants were themselves unaware, or indicated that the general public was unaware, that the tax credit being offered for EVs in Colorado was fully refundable. A fully refundable tax credit allows for EV purchasers to be fully reimbursed for the value of the tax credit regardless of their income tax. This mechanism eliminates the need for extensive tax liability to receive the tax credit benefits.

During meetings in Colorado, participants primarily listed the following factors as critical to EV adoption: **price, limited range, model availability,** and **limited information.** From discussions, I find it likely that dealers are not adequately informing potential EV adopters in Colorado about EVs and EV incentives. Potential buyers are further turned off of adoption because many dealers do not own EVs themselves, and have limited knowledge or experience with EVs and EV incentives. Several meeting participants stated that dealers would refer interested customers to tax accountants, rather than try to answer questions related to the EV income tax credit. Dealers indicated that these referrals were necessary because the incentive situation is unique to each individual, the tax credit eligibility differs, and that dealers themselves are not tax accountants and are not well versed in the tax code.

While support from the governor's office for EVs has been considerable, discussions in stakeholder meetings indicated that technology legitimation for potential adopters is still lacking in Colorado, and that widespread adoption of EVs in the greater Denver area has been limited given the strong income tax incentive. The deployment data through 2014 indicates that Chevy Volts have sold better than Leafs in most Colorado markets except for in Boulder, where Leaf adoption has been strong. Also interesting to note is that when questions were asked about EV leasing compared to outright purchases in Colorado, participants indicated that there was minimal adoption of EVs through leasing agreements. This is in stark contrast to other states, like Georgia, where participants stated that leasing arrangements have accounted for the majority of EV deployment in the state. Meeting participants also brought to my attention a new program in Boulder specific to the Boulder Nissan dealership – Solar Benefits Colorado.

Solar Benefits Colorado is a partnership between Sunrun, Nissan Boulder, and local communities to further accelerate deployment of solar panels in combination with Leafs by providing discounts based on "group" purchasing. Vote Solar orchestrated the arrangements for this discount program. For electric vehicles, the program brings the

MSRP for a new Leaf down to $23,461 from $31,810 prior to the federal and state tax credit, and further offers 24 months of complimentary charging. While marketing for this partnership has only just begun, stakeholders stated that widespread media attention has promoted considerable interest in Leafs in the Boulder area. Specifically, it was stated that once press about the program went out, interest in Leafs at the Boulder Nissan dealership sharply increased. In addition to the Sunrun-Nissan partnership, Nissan also offers very competitive 0% APR financing for up to three-years on newly purchased Leafs.

5.3 The Case of Georgia

During stakeholder meetings in Georgia, participants stated that EV penetration was primarily driven by two factors: **price** and **information**. Georgia, similar to Colorado, has offered an aggressive purchase incentive for electric vehicles. Unlike in Colorado, however, this tax incentive manifested in 1998 out of support for alternative fuel vehicles, such as natural gas vehicles, prior to the existence of modern EVs in the market. As newer model electric vehicles entered the market, they became eligible for a $5,000 state tax credit that worked to directly reduce income tax liability associated with EV purchases and leases for up to a 5-year period.

The Georgia income tax credit only applied to fully electric vehicles, excluding plug-in hybrid vehicles such as the Chevy Volt. Prior to the repeal of this incentive in July, 2015, the $5,000 income tax credit could be fully applied to the lease or purchase of a new electric vehicle. EV sales in Georgia have been significantly different from elsewhere in the country.

From model M5.1, I find that the state-year fixed effects for Georgia were highly significant in 2013 and 2014 at 95% and 99% confidence levels respectively. This indicates that EV sales in Georgia were significantly higher in 2013 and 2014 than could be explained by other variables, including the general environment that existed in Georgia in previous years. Meetings and newspaper article data were unable to say for certain why interest in EVs grew dramatically in 2013 and 2014 compared to prior years. Word-frequency analysis of newspaper articles, however, indicates that starting in 2013, articles started discussing EVs as a present technology, as opposed to a "future" technology. The most-frequently used word to describe EVs from 2008 through 2012 was "will." As in, "EVs *will* enter the market," or "Vehicles *will* be able to travel 100 miles on a single charge." Starting in 2013, however, articles started talking more about the technology in the present tense, with the top three most frequently used words for both years being "electric, car, and vehicle." Also important to note is that articles made frequent mention of Atlanta, while other metropolitan regions in Georgia were not frequently mentioned.

Figure 5.2. Word cloud map depicting the frequency of the most commonly used words in articles associated with the innovation system in Georgia.

In the United States, Nissan has been able to creatively work the $7,500 federal tax credit into their leasing plan, which enables a competitively priced lease rate of $199 per month for Nissan Leafs. Unique to the case of Georgia is that the $5,000 tax credit could be applied, in full, over a 3-year period of the lease agreement. This incentive is likely responsible for the sudden and aggressive deployment of Nissan Leafs in Georgia. As word got out about the ability to procure a three-year lease for a Nissan Leaf for approximately $2,000 total, EV uptake increased. Uptake was further aided by marketing, which portrayed driving EVs as "practically free." Stakeholders state that additional support for EV deployment has come from HOV lane access permissions, although not all stakeholders were aware of the costs or details of obtaining the decals to permit access to the HOV and high-occupancy toll lanes. Interview participants stated various costs associated with decal prices, and were not always certain how registration of decals occurred. Most meeting participants in Georgia owned electric vehicles.

Despite all the success that has occurred with EV deployment in Georgia, it appears that EV uptake has been limited compared to the level that we might expect. Given the very compelling economic incentive for EV uptake, we should see far greater adoption and utilization of EVs in the State. Instead, meeting participants indicated that EV adoption was mostly restricted to more affluent individuals in specific areas. When looking at the regional distribution of EVs, it becomes apparent that most EV deployment has occurred in Atlanta, as opposed to surrounding areas, such as Savannah (Figure 5.3).

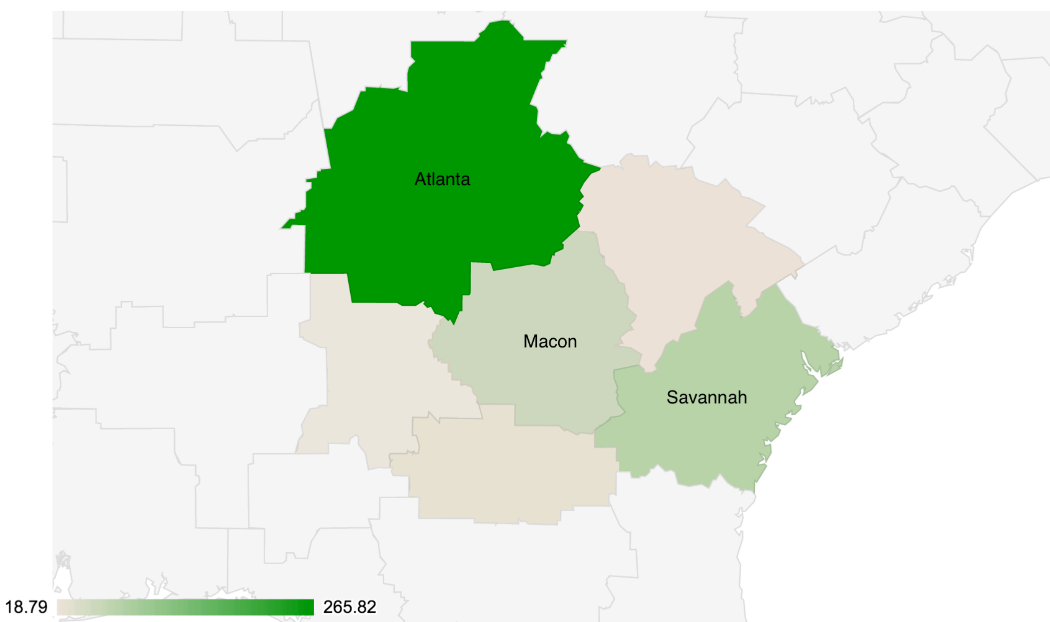

Figure 5.3. EV Deployment in Georgia per 100,000 people. Most EVs have been deployed in the Atlanta-metro area.

In meetings, participants mentioned the presence of a “champion” Nissan dealer in the Atlanta area. “Champion” dealers were not present elsewhere in the state. Despite the compelling economic incentives for EV adoption, limited diffusion in areas outside of Atlanta is indicative of limited knowledge diffusion between communities.

Also unique to Georgia is that the primary electric utility is actively engaging in EVSE deployment. Georgia Power and Nissan offer very compelling incentives for the installation of workplace charging. These incentives amount to $500 committed by Georgia Power with an additional $500 being matched by Nissan toward workplace EVSE. Other modest grant programs supported by state and local governments have aided the deployment of additional EVSE.

Meeting participants stated that since the income tax credit for EVs has been repealed, EV purchases have substantially declined. Recent R.L. Polk Data also shows a decline in EV sales since the tax credit has been repealed (Gordon-Bloomfield, 2015). In addition to the recent removal of this purchase incentive, the Georgia legislature has also adopted a road-use fee of $200 per year for registered electric vehicles. This is the largest EV road-use fee in the United States, and does not go toward supporting EVSE (the one in Washington does).

5.4 TIS Analysis

To provide better assessment of the EV innovation narrative in Washington, Colorado, and Georgia, I turn to the quantitative methods for Technology Innovation System analysis previously discussed in Chapters 2 and 4. This analysis relies on large bodies of textual information retrieved from newspapers, trade journals, magazines, and major blogs. Similar to how discussion questions were developed for in-person meetings to assess EV innovation, the data collected from these textual sources were classified and categorized into the 7 innovation functions.

Textual media relevant to electric vehicles has been captured for all 50 states (and D.C.) from 1994 through 2014. The content from each document has been classified as corresponding to one or more of the technology innovation system functions. The innovation system is then tracked over time. Support for each function, and how this support changes, can be displayed visually. Figure 5.4, Figure 5.6, and Figure 5.7 show the TIS mapping for Washington, Colorado, and Georgia in comparison to the overall TIS trend at the national level.

To derive meaning from this analysis, I employ statistical modeling techniques to assess if any innovation system trends are significantly different for Washington, Colorado, or Georgia compared to the nation as a whole. The following statistical model is used to assess each innovation system function (a negative binomial general linear model with a log linking function):

$$ln\{Classification\ Counts\}_{it} = \beta_t\{Year\}_t + \beta_r\{State\}_i + \beta_{rt}\{State{:}\,Year\}_{it} + \varepsilon_{it} + \alpha$$

This model assumes that state and time fixed effects account for a majority of the difference associated with media portrayal of electric vehicle innovation each month (t). The State:Year fixed effect accounts for differences that occur in a state in a given year. *If this effect is significant, it indicates that something unique happened in that state in that year compared to the expected trend*

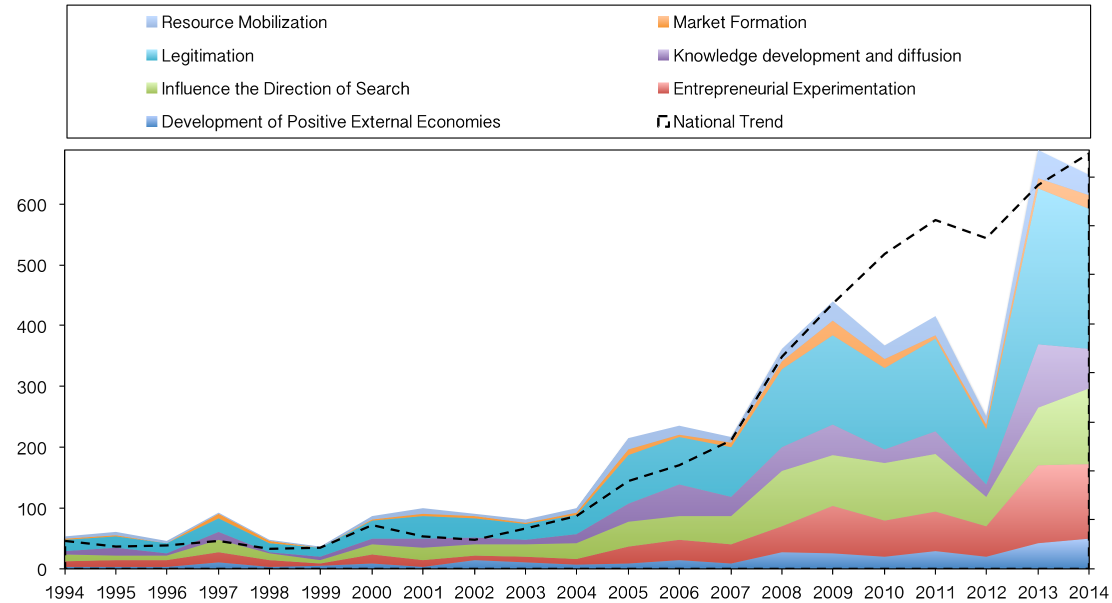

Figure 5.4. A visual mapping for Washington of the different attributes of the innovation system. The overall National trend is shown for comparison.

Modeling results indicate that Entrepreneurial Experimentation has been significantly better supported in Washington than elsewhere in the nation. This means that there is considerable interest in EVs from the private sector and private industry; this includes discussion of joint ventures, commercialization activity, business mergers and acquisitions, and launching of new products associated with electric vehicles. To better understand why the TIS declined in 2012, I examine word clouds for each year. Starting in 2012, a number of new EVs emerged (like the Tesla Model S), and so discussion of EVs suddenly shifted from dealing with a future technology ("EVs *will* be entering the market"), to a current technology ("The Tesla Model S is available"). It is therefore the large increase in EV models from 2012 to 2013, and the shift in media reporting from a future event to a current occurrence that causes the stark contrast between the 2012 and 2013 TIS in Washington.

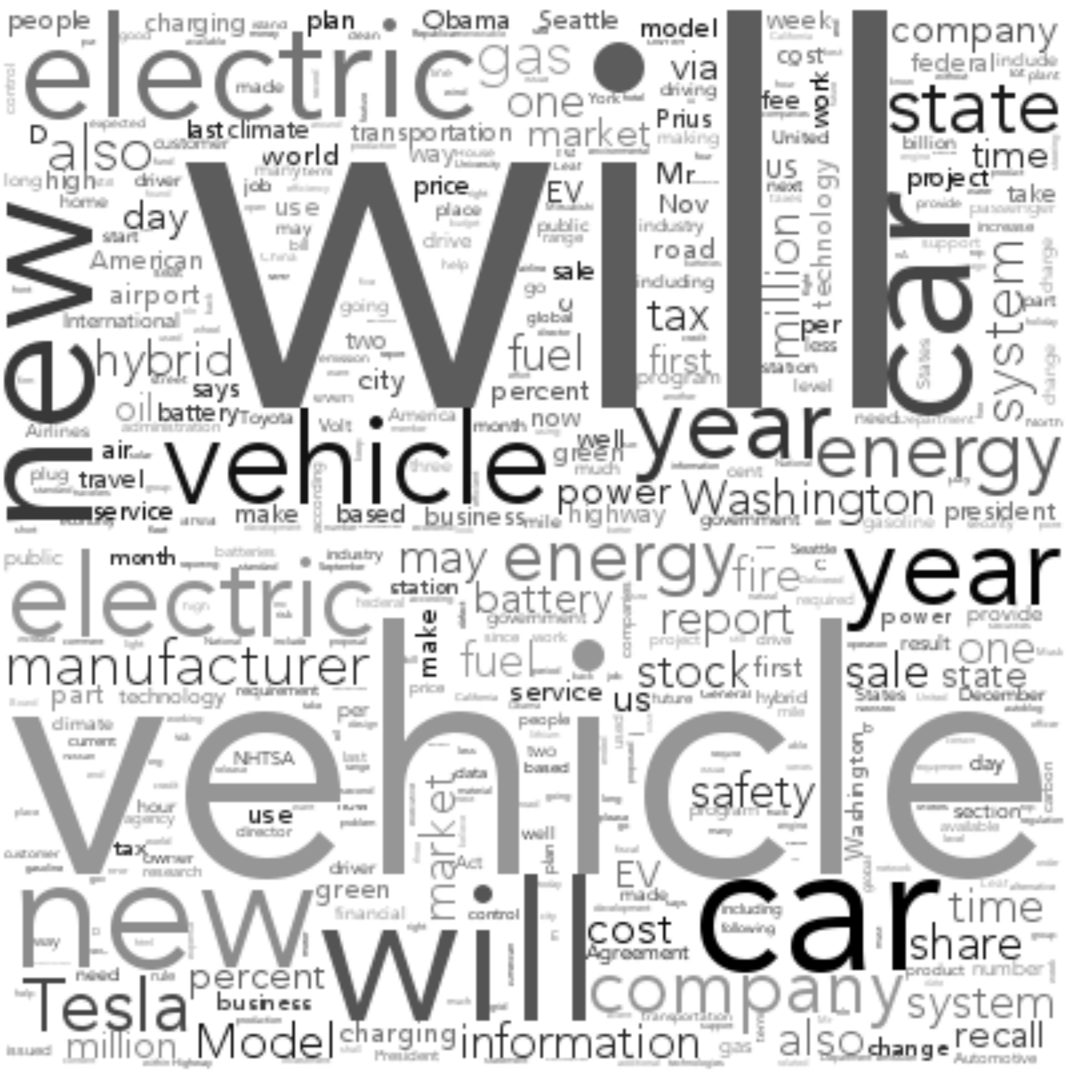

Figure 5.5. Word clouds for electric vehicles articles in Washington in 2012 (top), and in 2013 (bottom). Word usage indicates the technology began to transition from a “future” technology to a technology actively being used and deployed.

Comparatively, the Colorado-specific fixed effects are not significant when it comes to describing EV innovation. This indicates that *electrical vehicle innovation in Colorado is not significantly different from what is seen in other states*. This result aligns with the previous finding that EV deployment in Colorado has not been significantly different compared to elsewhere.

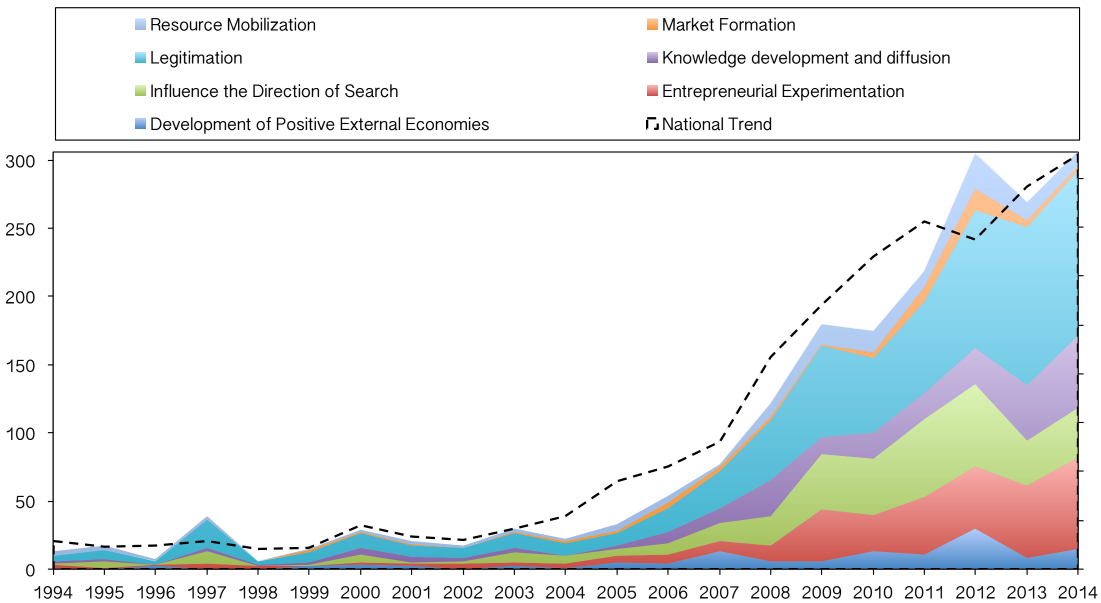

Figure 5.6. A visual mapping for Colorado of the different attributes of the innovation system. The overall National trend is shown for comparison.

When looking at the innovation system in Georgia, model results indicate that there are no positive shifts in the overall innovation system compared to elsewhere. However, model results also show *continuous and significant negative support in Georgia for the "Influence the Direction of Search" innovation function since 2010.* This negative relationship implies that statements of support or political action in Georgia have not directly signaled out or promoted the adoption of EVs relative to elsewhere. *This signal aligns with the recent repeal of the income tax credit that created economic incentives for EV adoption.*

One important thing to note from this analysis is that the “Resource Mobilization” function does not appear to be significant in Colorado or Georgia compared to elsewhere. This result is seemingly contradictory given that the income tax credits provided in Colorado and Georgia were much higher than the tax incentives and monetary supported offered elsewhere in the United States for EV adoption. However, if the price of EVs is not the fundamental driver of adoption in the state, or if the income tax credit has not been adequately taken advantage of at the state level, then I would not expect to see significant “Resource Mobilization” efforts compared to elsewhere. Although the resources may theoretically be there, they are not being properly mobilized or taken advantage of. In the case of Georgia, once EV sales became significant, resources that may have continued to support deployment were removed.

One weakness to this TIS approach is that it tracks only statewide trends, and is inadequate for assessing county-specific of city-specific outcomes. As was discussed previously, most of the innovation activities and resource mobilization efforts in Georgia have predominately manifested in Atlanta, where knowledge was disseminated throughout local communities. This means that the rest of the state likely has not been supportive of EVs, which may further explain recent antagonism toward EV deployment in the legislature.

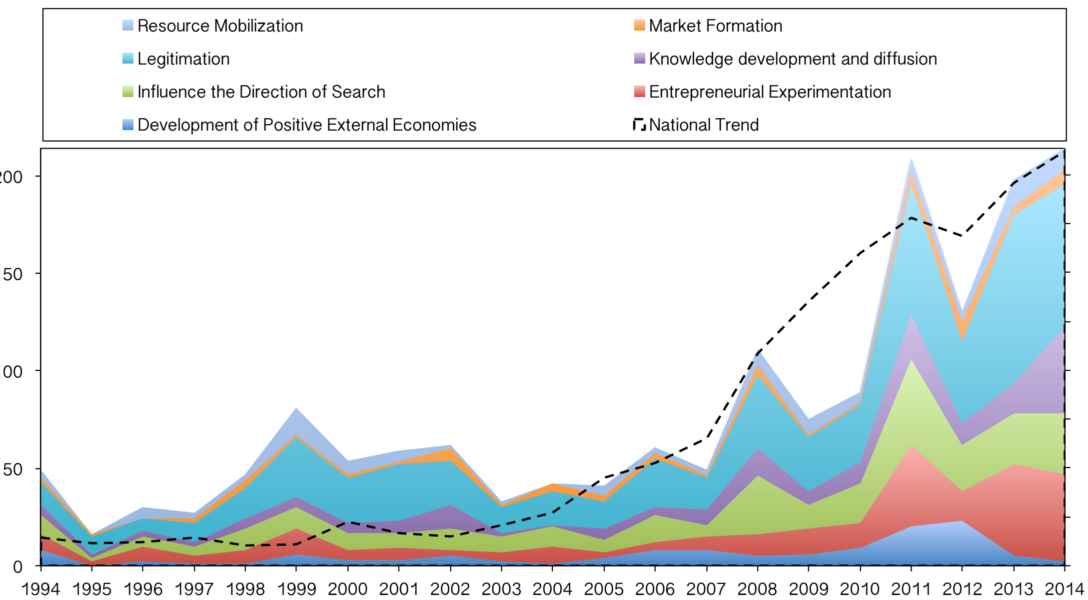

Figure 5.7. A visual mapping for Georgia of the different attributes of the innovation system. The overall National trend is shown for comparison.

Quantitative analysis of TIS trends can also be used to facilitate discussion and understanding about what might be done to improve electric vehicle uptake in the future. To arrive at a set of suggestions, I look at the qualitative data gathered from in-person meetings in addition to textual data analysis and statistical modeling. Combined, these approaches can facilitate and ground the discussion on steps forward for accelerating EV deployment.

To better explore EV adoption trends, I use the textual data I collected and classified as the only explanatory variable for EV deployment in a general linear model. I utilize a statistical model (a negative binomial general linear model with a log linking function) to predict electric vehicle deployment based on the 7 different innovation functions without additional explanatory variables. The following model formulation is used for each innovation function:

$$ln\{Vehicles\ Sold\}_{it} = \beta_1\{State\}_i + \beta_{2\ldots8}\{C_{IF}\}_{it} + \varepsilon_{it} + \alpha$$

where C_{IF} is the count of textual documents in the database that were classified to correspond to a given innovation functions (IF) for region i and month t, ε is the error term, and α is the intercept. Regression results from this model are shown in Table 5.1. I compare the TIS model to a Naïve forecast of EV uptake. The Naïve forecast is a simplistic extrapolation of past trends to provide a comparable case for what EV deployment might be. For the Naïve forecast, I only consider a linear time trend and state-based fixed effect, serving to change the intercept of the linear time trend:

$$ln\{Vehicles\ Sold\}_{\mathrm{it}} = \beta_1\{State\}_i + \beta_2\{time\}_t + \varepsilon_{it} + \alpha$$

The model fit for the Naïve forecast compared to the TIS article count model for each of the three states considered is shown below in Figure 5.8, Figure 5.9, and Figure 5.10.

The TIS model for Washington has an R^2 value of 0.27, indicating that the TIS system functions can explain 27% of EV deployment in the state over time. The Naïve model trends exponential, despite the linear time trend, because of the negative binomial model specification. For negative binomial models, a logarithmic link function is used. The logarithmic link function is necessary for this data because EV sales have to be greater than or equal to 0 for any given monthly period. Negative binomial models are often used to assess count data, and to account for any 0 counts that may occur for the observed sales trend, and to account for overdispersion or differences compared to a typical Poisson distribution.

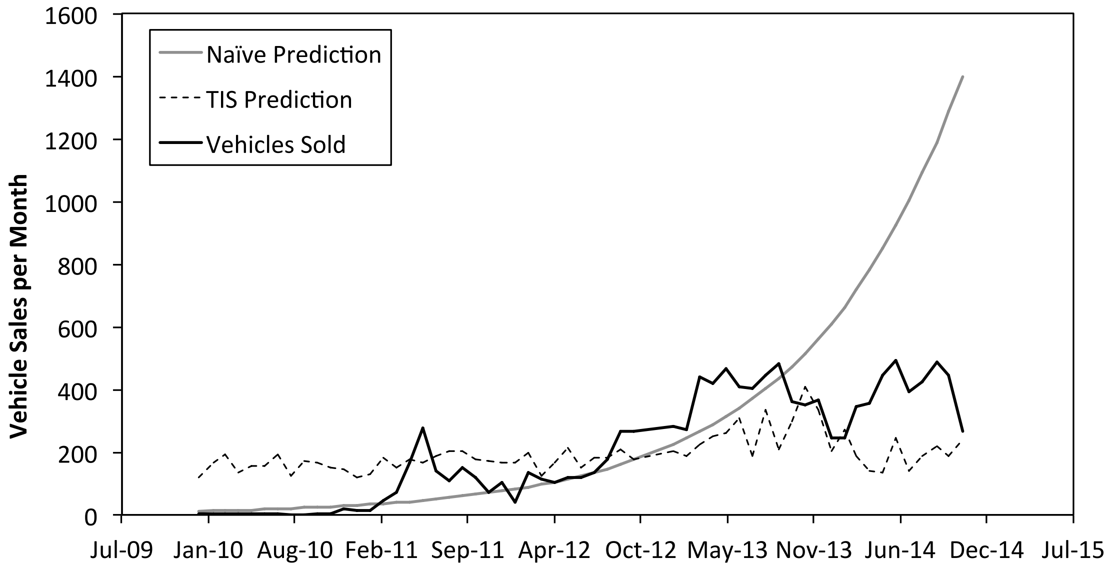

Figure 5.8. Predicted EV sales compared to actual EV sales in Washington. TIS Prediction has an R^2 value of 0.27

As seen in Figure 5.8, TIS model prediction is more inline with EV sales in Washington than the Naïve forecast model. For Colorado, a similar result is shown. The TIS model for Colorado has an R^2 value of 0.20 – that is to say, the TIS model explains 20% more of the trend for EV sales than a trend line of average monthly sales. The TIS prediction model once again appears to provide more explanation of EV deployment in Colorado than the Naïve forecast model.

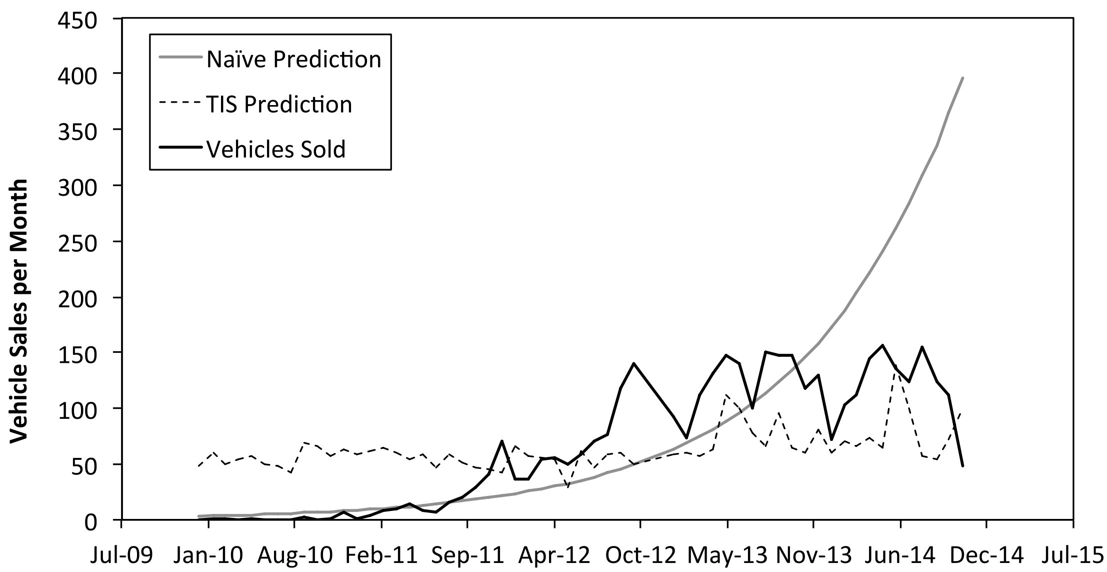

Figure 5.9. Predicted EV sales compared to actual EV sales in Colorado. TIS Prediction has an R^2 value of 0.20

For Georgia, the TIS prediction model no longer fits the data better than the Naïve forecast. The TIS model has an R^2 value of only 0.06 – implying weak explanatory power of the trends occurring in Georgia. Because the Naïve forecast appears to fit the data, it is possible that vehicle uptake in Georgia was limited by information diffusion, as opposed to other barriers or factors. Information and technology diffusion is often represented by the Bass diffusion model which predicts deployment as an exponential or logistic function over time (Dodson & Muller, 1978; Geroski, 2000). A typical, knowledge-based diffusion model would appear similar to the diffusion trend shown for EV sales in Georgia.

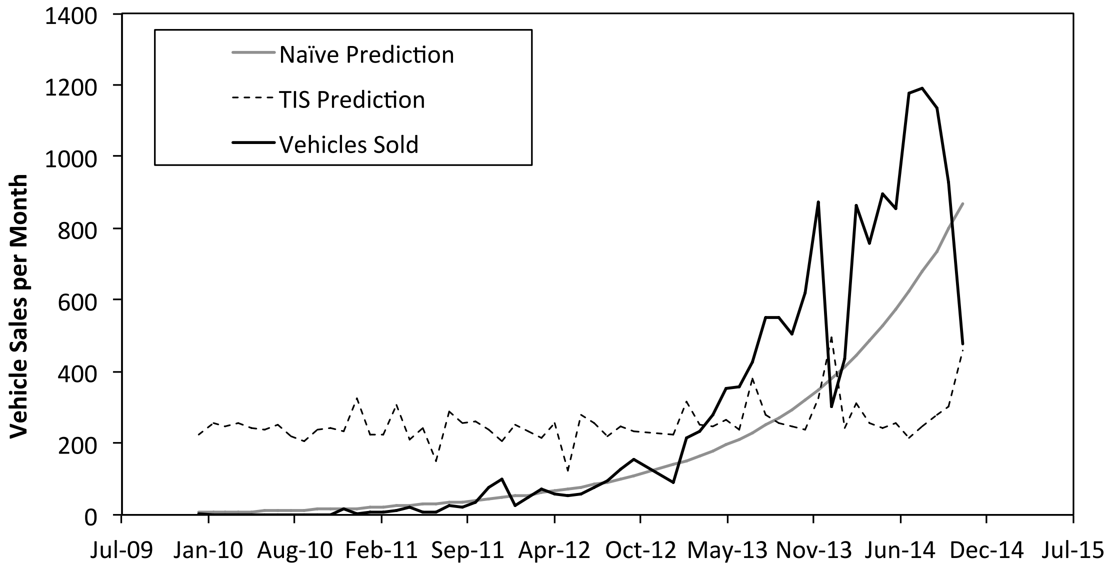

Figure 5.10. Predicted EV sales compared to actual EV sales in Georgia. TIS Prediction has an R^2 value of 0.06

Although the TIS model can only explain around 20% of the observed vehicle sales trends for Colorado and Washington, it can nevertheless be used to explore the innovation topics most relevant for predicting EV deployment. The modeling results shown in Table 5.1 indicate that technology "legitimation" and "knowledge development and diffusion" are highly significant, positive predictors of EV deployment. The results from this analysis support the results and information collected through meetings and discussions with key stakeholders, mainly that information and knowledge dissemination for EVs has been limited in each state, and that information is a critical factor in driving EV deployment.

From these independent methods of analysis, it becomes evident that informational outreach and education can play a significant role in promoting EV uptake, and that more effort on these fronts is likely necessary to promote EV adoption in the future. Meetings and discussions revealed that there is a fundamental knowledge gap for dealers and for consumers when it comes to EV incentives and familiarity with the technology. Similarly, in Colorado the perception that 4WD vehicles provide increased safety or improved handling in icy or snowy conditions is another barrier to EV adoption that could be better addressed through education and knowledge diffusion efforts.

Dealerships play an essential role in disseminating information and knowledge to potential customers and in legitimizing the technology through test drives and familiarity (Cahill, Davies-Shawhyde, & Turrtentine, 2014). Without properly educated dealers, customers that may be a good fit for EV ownership are neither being steered toward potential EV choices, nor are their concerns or existing questions about EV technology being answered at this point of interaction.

From meetings with key representatives, a narrative emerged in which EV uptake was furthered when customers were incentivized to lease EVs compared to making outright purchases. Both Washington and Georgia show increased diffusion of EVs compared to Colorado where leasing is not as prevalent. Leases may further help legitimize the technology by facilitating longer trial periods for consumers to become familiar with the technology, and lowers the barrier to adoption. This not only improves the technology legitimacy, but may also serve to increase diffusion of knowledge and information about EVs. There is some evidence for this in the EV analysis conducted for Georgia and Washington.

Table 5.1

Regression results for article counts as a predictor of monthly EV sales

	Dependent variable:
	Vehicles_Sold
Legitimation	0.024***
	(0.006)
InfD	-0.046***
	(0.009)
MarketFormation	-0.019
	(0.024)
EntExp	0.010
	(0.010)
KnowD	0.049***
	(0.009)
PosEx	-0.034**
	(0.013)
ResMob	-0.023
	(0.014)
StateCALIFORNIA	1.705***
	(0.496)
StateCOLORADO	-0.119
	(0.242)
StateCONNECTICUT	-0.571**
	(0.242)
StateFLORIDA	0.854***
	(0.243)
StateGEORGIA	1.319***
	(0.241)
StateILLINOIS	0.380
	(0.242)
StateINDIANA	-0.848***
	(0.243)
StateMAINE	-1.787***

	(0.245)
StateMARYLAND	0.207
	(0.241)
StateMASSACHUSETTS	-0.064
	(0.244)
StateMICHIGAN	0.517*
	(0.265)
StateMINNESOTA	-0.560**
	(0.242)
StateNEW JERSEY	0.293
	(0.241)
StateNEW MEXICO	-1.883***
	(0.245)
StateNEW YORK	0.012
	(0.282)
StateNORTH CAROLINA	-0.251
	(0.242)
StateOREGON	0.309
	(0.241)
StateRHODE ISLAND	-2.351***
	(0.247)
StateTENNESSEE	-0.434*
	(0.242)
StateTEXAS	0.685***
	(0.244)
StateVERMONT	-1.604***
	(0.244)
StateVIRGINIA	-0.155
	(0.242)
StateWASHINGTON	0.985***
	(0.243)
Constant	4.148***
	(0.172)
Observations	1,368

Log Likelihood	-6,986.192
theta	0.608^{***} (0.023)
Akaike Inf. Crit.	14,034.380
Note:	*p<0.1; **p<0.05; ***p<0.01

Chapter 6: Conclusions

This dissertation examines technology innovation, and investigates methodologies for assessing past technology innovation trends for biofuels and electric vehicles. These trends help reveal what has or hasn't happened for diffusing and innovating on these technologies. The insights, in-turn, can lead to meaningful discussion and plans of action for stakeholders and policy makers moving forward.

In Chapter 3, I examined the use of patents for assessing technology innovation. Patents have traditionally been used as a proxy for innovation, which makes patent data an important starting point for comparisons. My analysis shows that the innovation narratives supported by patent trends is substantially different depending on the classification scheme used. Through literature review and discussion with biofuel experts, a default biofuel innovation narrative was established.

Two traditional patent classification schemes were used to compare this literature-based narrative to the patent data. Both the International Patent Classification system and the newer Cooperative Patent Classification system were used. These two classification systems provided contradictory narrative results, in addition to having different statistical relationships with explanatory variables. The Natural Language Processing approach that I pioneered for this dissertation better aligned with the literature-based narrative, and yielded intuitive statistical relationships.

Between the literature narrative and the NLP classification results, several insights about the biofuel innovation process have been gained. For instance, data support that 1st generation biofuel innovation in the U.S. has likely benefited from spillover effects from genetically engineered food crops. Data also support the idea that the shift in agriculture to

GE crops laid the groundwork for the rapid expansion of the 1st generation biofuel industry starting in the '90s.

All narratives and methods employed to assess biofuel innovation show that after the Renewable Fuel Standard was adopted in 2005, a new period of innovation began. This innovation period may have encouraged new firms to experiment with 2nd generation biofuels. Most new firms filing for patents after the enactment of the RFS were primarily firms associated with biotechnology, rather than agricultural, but the number of 2nd generation biofuel patents by both types of companies was small. Patent analysis indicated that this shortfall in 2nd generation biofuel innovation was not been filled by the oil and gas industry.

For the timespan that my data cover, there is indication that commercial investment in 2nd generation facilities has been limited. This suggests the need for new R&D (and investment) policy approaches if major expansion of 2nd generation biofuel is desired.

In addition to using patent data and the biofuel literature narrative, I employed a novel dataset to better provide insight into the innovation process. In Chapter 4, I employ the use of an informational dataset that I created which is grounded in Technology Innovation System (TIS) theory.

The TIS-approach to innovation recognizes that flows of information are quintessential for promoting successful innovation outcomes. The TIS uniquely classifies these information flows into 7 important functions, or types of information associated with specific actions that can be taken to promote innovation. I utilized computational techniques (natural language processing) in conjunction with the large textual database of newspapers, trade journals, and other text-based news sources to establish temporal trends for the different types of information flow for biofuels. In Chapter 5 I further applied this methodology to assess electric vehicle innovation trends. I show that when using an information-based approach to innovation (using text-media), it is possible to not only predict patent trends, but that article counts also correspond to policy implementation, and that article data may further align with technology deployment trends.

Similar to patent data, the Biofuel TIS data show the fundamental shift in innovation that occurred following enactment of the Renewable Fuel Standard. Importantly, however, these data also show that simply enacting the RFS is not the whole story, and that there is nuance associated with how the policy has morphed over time, and how flows of information have changed.

Also, the TIS data can be used to assess state-level policy, something difficult to do with patents. Working from the Alternative Fuel Data Center's policy database, I was able to use TIS data in conjunction with policy enactment dates to flush out an innovation narrative for California, and to better evaluate the effect that the California policy landscape has had on fostering biofuel innovation in the State. Overall, this approach can be used to facilitate direct discussion with policy makers to better understand the actions that have been taken, or that could be taken in the future, to direct or promote innovation or deployment goals.

In Chapter 5, I extensively relied on the TIS-framework to ground and direct discussions with stakeholders across three different states. From both stakeholder discussions and through the NLP-based informational dataset, it became apparent that a major deterrent for EV adoption was a gap in knowledge and information for EVs.

A follow-up discussion with EV stakeholders in Colorado similarly supported this conclusion. Drive Electric Northern Colorado, for instance, indicated that they have started to take a more active approach in educational outreach for 2015. Although anecdotal, Drive Electric indicates that since outreach began, they have seen a tremendous uptick in sales in the Northern Colorado compared to what had been expected.

These knowledge gaps for EVs aren't fundamental technology gaps or deficiencies, but are instead gaps pertaining to institutional factors such as purchase incentives, and supporting infrastructure. These are key elements of the TIS framework that are not captured by traditional innovation models. While the Colorado case of the AWD culture may exemplify a technological deterrent from EV adoption, the fact that this stigma has been broken in Boulder County, but not in Denver, more clearly epitomizes the lack of knowledge diffusion to potential car buyers. This knowledge gap is also apparent in Georgia, where EV sales are once again higher in areas near "champion" dealers, despite the compelling economic case for EV ownership that existed for almost all commuters.

Using the TIS analysis approaches extensively detailed in this dissertation, these knowledge gaps and trends become clear, while they would have remained obfuscated using more traditional approaches for measuring, assessing, and predicting innovation. The information gleaned from the innovation analysis approaches used in this dissertation can further be used to ground discussion and to foster strategies moving forward that may best work to encourage technology adoption and innovation.

6.1 Future Work

While this dissertation establishes a set of new methods to aid policy makers in better understanding technology innovation, additional research is still necessary. Specifically, there were fundamental gaps in the data used for this dissertation – especially for the vehicles sales data and other technology deployment data, which were difficult to obtain, especially in more disaggregate forms.

To better establish and assess the methodologies pioneered in this dissertation, I suggest that further work be done to examine historic innovation cases, especially for energy transitions. Given the relative newness of the technologies examined here, the time series were too small to facilitate many traditional and more rigorous statistical assessments. Historical newspaper archives alongside historical innovation cases can be used to better assess the legitimacy of the methods I employed.

Furthermore, there is need to create a larger, more robust machine learning dataset to be used for classifying articles for other technologies, not just biofuels and electric vehicles. As more technologies are added, and additional manual classification occurs, it may be possible to better parse out the defining characteristics of the 7 innovation functions. Such an endeavor could better facilitate the parameterization of the TIS framework, which could aid in predicting technology outcomes. Parameterization would allow the TIS-approach to innovation to be incorporated into traditional modeling techniques as well as integrated assessment models. This would make it possible to better inform economic models and engineering models on likely outcomes of policy implementation.

Still missing in this work is the theoretical basis for how specific innovation functions ultimately impact one another, leading to positive feedback loops and cascade mechanisms. It may be possible to better integrate some of the data collected here into system dynamic models to further explore the relationships between different innovation functions and innovation outcomes. Using similar data-heavy approaches it may ultimately be possible to break into the innovation black box, providing a more detailed, nuanced model that can more accurately predict and show how scarce resources should be allocated to promote future innovations.

References

64th Washington Legislature. (2015). *Second engrossed substitute senate bill 5987*.

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy, 31*(7), 1069-1085. doi: 10.1016/s0048-7333(01)00184-6

Adams, P., Fontana, R., & Malerba, F. (2013). The magnitude of innovation by demand in a sectoral system: The role of industrial users in semiconductors. *Research Policy, 42*(1), 1-14. doi: 10.1016/j.respol.2012.05.011

Adelman, D. E., & Deanglis, K. L. (2007). Patent metrics: the mismeasure of innovation in the biotech patent debate. *Texas Law Review, 85*, 1677.

Agrawal, A., & Henderson, R. (2002). Putting Patents in Context: Exploring Knowledge Transfer from MIT. *Management Science, 48*(1), 44-60. doi: 10.1287/mnsc.48.1.44.14279

Alberth, S. (2008). Forecasting technology costs via the experience curve—Myth or magic? *Technological Forecasting and Social Change, 75*(7), 952-983.

Allison, J. R., & Lemley, M. A. (2000). Who's Patenting What-An Empirical Exploration of Patent Prosecution. *Vand. L. Rev., 53*, 2099.

Archibugi, D. (1992). Patenting as an indicator of technological innovation: a review. *Science and Public Policy, 19*(6), 357-368.

Arrow, K. (1962). Economic welfare and the allocation of resources for invention *The rate and direction of inventive activity: Economic and social factors* (pp. 609-626): Princeton University Press.

Asur, S., & Huberman, B. A. (2010). Predicting the Future with Social Media. 492-499. doi: 10.1109/wi-iat.2010.63

Audretsch, D. B. (2007). Entrepreneurship capital and economic growth. *Oxford Review of Economic Policy, 23*(1), 63-78. doi: 10.1093/oxrep/grm001

Balan, V., Chiaramonti, D., & Kumar, S. (2013). Review of US and EU initiatives toward development, demonstration, and commercialization of lignocellulosic biofuels. *Biofuels, bioproducts and biorefining, 7*(6), 732-759. doi: 10.1002/bbb.1436

Beniger, J. R. (1986). *The control revolution: Technological and economic origins of the information society*: Harvard University Press.

Bergek, A., & Jacobsson, S. (2003). The emergence of a growth industry: a comparative analysis of the German, Dutch and Swedish wind turbine industries *Change, Transformation and Development* (pp. 197-227): Physica-Verlag HD.

Bergek, A., Jacobsson, S., Carlsson, B., Lindmark, S., & Rickne, A. (2008). Analyzing the functional dynamics of technological innovation systems: A scheme of analysis. *Research Policy, 37*(3), 407-429. doi: 10.1016/j.respol.2007.12.003

Boone, C., & Ozcan, S. (2013). Why Do Cooperatives Emerge in a World Dominated by Corporations? The Diffusion of Cooperatives in the U.S. Bio-Ethanol Industry, 1978-2013. *Academy of Management Journal, 57*(4), 990-1012. doi: 10.5465/amj.2012.0194

Brekke, K. (2010). Ethanol in California: Past Success, Future Questions.

Breschi, S. (2001). Knowledge Spillovers and Local Innovation Systems: A Critical Survey. *Industrial and Corporate Change, 10*(4), 975-1005. doi: 10.1093/icc/10.4.975

Breschi, S., Malerba, F., & Orsenigo, L. (2000). Technological Regimes and Schumpeterian Patterns of Innovation. *The Economic Journal, 110*(463), 388-410. doi: 10.1111/1468-0297.00530

Brown, C., & Linden, G. (2009). *Chips and change: how crisis reshapes the semiconductor industry*: MIT Press.

Brown, R., Collantes, G., Eggert, A., Handy, S., Kessler, J., Shulock, C., . . . Yeh, S. (2015). Achieving California's Greenhouse Gas Goals: A Focus on Transportation (E. a. t. E. The Policy Institute for Energy, UC Davis, Trans.). Next 10.

Cahill, E. C., Davies-Shawhyde, J., & Turrtentine, T. S. (2014). New Car Dealers and Retail Innovation in California's Plug-In Electric Vehicle Market.

California Air Resources Board. (2011). Low Carbon Fuel Standard 2011 Program Review Report (pp. 189).

Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine, 9*(2), 48-57.

Cameron, A. C., & Trivedi, P. K. (1986). Econometric models based on count data. Comparisons and applications of some estimators and tests. *Journal of applied econometrics, 1*(1), 29-53.

Carlsson, B., & Eliasson, G. (1995). *The nature and importance of economic competence*: Springer.

Carlsson, B., & Stankiewicz, R. (1991). On the nature, function and composition of technological systems. *Journal of Evolutionary Economics, 1*(2), 93-118. doi: 10.1007/bf01224915

Carriquiry, M. (2007). US biodiesel production: recent developments and prospects. *Iowa Ag Review, 13*(2), 8-9, 11.

Carriquiry, M. A., Du, X., & Timilsina, G. R. (2011). Second generation biofuels: Economics and policies. *Energy Policy, 39*(7), 4222-4234. doi: 10.1016/j.enpol.2011.04.036

Clarke, D., Jablonski, S., Moran, B., Anandarajah, G., & Taylor, G. (2009). How can accelerated development of bioenergy contribute to the future UK energy mix? Insights from a MARKAL modelling exercise. *Biotechnology for biofuels, 2*(1), 13.

Coenen, L., & Díaz López, F. J. (2010). Comparing systems approaches to innovation and technological change for sustainable and competitive economies: an explorative study into conceptual commonalities, differences and complementarities. *Journal of Cleaner Production, 18*(12), 1149-1160. doi: 10.1016/j.jclepro.2010.04.003

Collantes, G. (2010). Do green tech policies need to pass the consumer test?: The case of ethanol fuel. *Energy Economics, 32*(6), 1235-1244.

Collantes, G., & Sperling, D. (2008). The origin of California's zero emission vehicle mandate. *Transportation Research Part A: Policy and Practice, 42*(10), 1302-1313.

Corts, K. S. (2010). Building out alternative fuel retail infrastructure: Government fleet spillovers in E85. *Journal of Environmental Economics and Management, 59*(3), 219-234. doi: 10.1016/j.jeem.2009.09.001

Costantini, V., Crespi, F., & Curci, Y. (2013). BioPat: An Investigation Tool for Analysis of Industry Evolution, Technological Paths and Policy Impact in the Biofuels Sector. In V. Costantini & M. Mazzanti (Eds.), *The Dynamics of Environmental and Economic Systems* (pp. 203-226): Springer Netherlands.

Dill, G. M. (2005). Glyphosate-resistant crops: history, status and future. *Pest Manag Sci, 61*(3), 219-224. doi: 10.1002/ps.1008

Dodson, J. A., & Muller, E. (1978). Models of New Product Diffusion Through Advertising and Word-of-Mouth. *Management Science, 24*(15), 1568-1578.

Eggert, H., & Greaker, M. (2014). Promoting Second Generation Biofuels: Does the First Generation Pave the Road? *Energies, 7*(7), 4430-4445. doi: 10.3390/en7074430

Energy Information Administration. (2015). *Data 1: Motor Gasoline Sales Through Retail Outlets Prices* Retrieved from: http://www.eia.gov/dnav/pet/pet_pri_allmg_a_EPM0_PTC_Dpgal_m.htm

Environmental Protection Agency. (2015). FuelEconomy.gov. 2015, from http://www.fueleconomy.gov/feg/Find.do?action=sbs&id=37066&id=33558&id=30979

Ernst, D. (2005). Limits to Modularity: Reflections on Recent Developments in Chip Design. *Industry & Innovation, 12*(3), 303-335. doi: 10.1080/13662710500195918

Fagerberg, J. (2003). Innovation: A Guide to the Literature: Centre for Technology, Innovation and Culture, University of Oslo.

Fausti, S. W. (2015). The causes and unintended consequences of a paradigm shift in corn production practices. *Environmental Science & Policy, 52*, 41-50. doi: 10.1016/j.envsci.2015.04.017

Fernandez-Cornejo, J. (2009). *First decade of genetically engineered crops in the United States*: DIANE Publishing.

Fernandez-Cornejo, J., Wechsler, S., Livingston, M., & Mitchell, L. (2014). Genetically Engineered Crops in the United States. *US Department of Agriculture Economic Research Service*.

Foody, G. M., & Mathur, A. (2004). A relative evaluation of multiclass image classification by support vector machines. *IEEE Transactions on Geoscience and Remote Sensing, 42*(6), 1335-1343. doi: 10.1109/tgrs.2004.827257

Fouquet, R., & Pearson, P. J. G. (2012). Past and prospective energy transitions: Insights from history. *Energy Policy, 50*, 1-7. doi: 10.1016/j.enpol.2012.08.014

Franco, G., Cayan, D., Luers, A., Hanemann, M., & Croes, B. (2007). Linking climate change science with policy in California. *Climatic Change, 87*(S1), 7-20. doi: 10.1007/s10584-007-9359-8

Fulton, L., Morrison, G. M., Parker, N. C., Witcover, J., & Sperling, D. (2014). NextSTEPS White Paper: Three Routes Forward – Incremental, Transitional, and Leapfrog.: Institute of Transportation Studies.

G Cassman, K., & Liska, A. J. (2007). Food and fuel for all: realistic or foolish? *Biofuels, bioproducts and biorefining, 1*(1), 18-23.

Gårder, P. E. (2014). *All-Wheel Drive and Winter Weather Safety.* Paper presented at the Transportation Research Board 93rd Annual Meeting.

Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. *Research Policy, 31*(8-9), 1257-1274. doi: 10.1016/s0048-7333(02)00062-8

Geroski, P. A. (2000). Models of technology diffusion. *Research Policy, 29*(4-5), 603-625. doi: 10.1016/s0048-7333(99)00092-x

Glimstedt, H., Bratt, D., & Karlsson, M. P. (2010). The decision to make or buy a critical technology: semiconductors at Ericsson, 1980-2010. *Industrial and Corporate Change, 19*(2), 431-464. doi: 10.1093/icc/dtq011

Gordon-Bloomfield, N. (2015). With Electric Car Incentives Gone, Electric Car Registrations in State of Georgia Plummet. Retrieved 11/8/2015, 2015, from https://transportevolved.com/2015/11/03/with-electric-car-incentives-gone-electric-car-registrations-in-state-of-georgia-plummet/

Griliches, Z. (1998). Patent statistics as economic indicators: a survey *R&D and productivity: the econometric evidence* (pp. 287-343): University of Chicago Press.

Griliches, Z., Pakes, A., & Hall, B. H. (1988). The value of patents as indicators of inventive activity: National Bureau of Economic Research Cambridge, Mass., USA.

Grimmer, J., & Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis, 21*(3), 267-297. doi: 10.1093/pan/mps028

Grubler, A. (2012). Energy transitions research: Insights and cautionary tales. *Energy Policy, 50*, 8-16. doi: 10.1016/j.enpol.2012.02.070

Grübler, A., Nakićenović, N., & Victor, D. G. (1999). Dynamics of energy technologies and global change. *Energy Policy, 27*(5), 247-280. doi: 10.1016/s0301-4215(98)00067-6

Gül, T., Kypreos, S., & Barreto, L. (2007). *Hydrogen and biofuels–A modeling analysis of competing energy carriers for western Europe.* Paper presented at the Proceedings of the World Energy Congress—Energy Future in an Interdependent World||, Rome, Italy.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market Value and Patent Citations. *The RAND Journal of Economics, 36*(1), 16-38. doi: 10.2307/1593752

Hanemann, W. M. (2007). How California Came to Pass AB 32, the Global Warming Solutions Act of 2006. *Department of Agricultural & Resource Economics, UCB*.

Hekkert, M. P., Suurs, R. A. A., Negro, S. O., Kuhlmann, S., & Smits, R. E. H. M. (2007). Functions of innovation systems: A new approach for analysing technological change. *Technological Forecasting and Social Change, 74*(4), 413-432. doi: 10.1016/j.techfore.2006.03.002

Heller, M. A. (1998). Can Patents Deter Innovation? The Anticommons in Biomedical Research. *Science, 280*(5364), 698-701. doi: 10.1126/science.280.5364.698

Herbert, B., Szarvas, G., & Gurevych, I. (2010). Prior Art Search Using International Patent Classification Codes and All-Claims-Queries. *6241*, 452-459. doi: 10.1007/978-3-642-15754-7_54

Hopkins, D. J., & King, G. (2010). A Method of Automated Nonparametric Content Analysis for Social Science. *American Journal of Political Science, 54*(1), 229-247. doi: 10.1111/j.1540-5907.2009.00428.x

Horsky, D., & Simon, L. S. (1983). Advertising and the Diffusion of New Products. *Marketing Science, 2*(1), 1-17. doi: 10.1287/mksc.2.1.1

Howell, S. (1997). US Biodiesel Standards-An Update of Current Activities: SAE Technical Paper.

Hurvich, C. M., & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika, 76*(2), 297-307. doi: 10.1093/biomet/76.2.297

IEA. (2000). *Experience curves for energy technology policy*. Paris: International Energy Agency.

IPCC. (2011). Summary for Policymakers. Figure SPM.6. In R. P. M. O. Edenhofer, Y.; Sokona, K.; Seyboth, P.; Matschoss, S.; Kadner, & P. E. T.; Zwickel, G. Hansen; S. Schlomer; C. von Stechow (eds) (Eds.), *IPCC Special Report on Renewable Energy Sources and Climate*

Change Mitigation. Cambridge, United Kingdom and New York, NY, USA.

Jamasb, T. (2007). Technical change theory and learning curves: patterns of progress in electricity generation technologies. *The Energy Journal*, 51-71.

Janda, K., Kristoufek, L., & Zilberman, D. (2012). Biofuels: Policies and impacts. *Agricultural Economics, 58*(8), 372-386.

Jenkins, J., Muro, M., Nordhaus, T., Shellenberger, M., Tawney, L., & Trembath, A. (2012). Beyond boom and bust: Putting clean tech on a path to subsidy independence. *Breakthrough institute, Brookings institution, Massachusetts*.

Jenkins, J. D. (2014). Political economy constraints on carbon pricing policies: What are the implications for economic efficiency, environmental efficacy, and climate policy design? *Energy Policy, 69*, 467-477. doi: 10.1016/j.enpol.2014.02.003

Johnson, A. (2001). *Functions in innovation system approaches.* Paper presented at the Paper for DRUID's Nelson-Winter Conference.

Judson, R. A., & Owen, A. L. (1999). Estimating dynamic panel data models: a guide for macroeconomists. *Economics letters, 65*(1), 9-15.

Kalish, S. (1985). A New Product Adoption Model with Price, Advertising, and Uncertainty. *Management Science, 31*(12), 1569-1585. doi: 10.1287/mnsc.31.12.1569

Karshenas, M. D., & Stoneman, P. (1992). A flexible model of technological diffusion incorporating economic factors with an application to the spread of colour television ownership in the UK. *Journal of Forecasting, 11*(7), 577-601. doi: 10.1002/for.3980110702

Keefe, R., Griffin, J. P., & Graham, J. D. (2008). The benefits and costs of new fuels and engines for light ‐ duty vehicles in the United States. *Risk Analysis, 28*(5), 1141-1154.

Kelly, F., & Brannon, J. I. (1996). ISTEA: another subsidy for alternative fuels. *Transportation Quarterly, 50*(2).

Kendall, A., & Yuan, J. (2013). Comparing life cycle assessments of different biofuel options. *Curr Opin Chem Biol, 17*(3), 439-443. doi: 10.1016/j.cbpa.2013.02.020

Kessler, J. (2015a). Biofuel Article Classifications (Publication no. doi/10.7910/DVN/KQIUBZ). from Harvard Dataverse http://dx.doi.org/10.7910/DVN/KQIUBZ

Kessler, J. (2015b). EV TIS Classifiers (Publication no. doi/10.7910/DVN/ZXB6BA). from Harvard Dataverse http://dx.doi.org/10.7910/DVN/ZXB6BA

Kessler, J. (2015c). *Stanford NLP Model Output for Biofuel Patent Classification*. Retrieved from: http://dx.doi.org/10.7910/DVN/29374

Kim, J., & Marschke, G. (2004). Accounting for the recent surge in U.S. patenting: changes in R&D expenditures, patent yields, and the high tech sector. *Economics of Innovation and New Technology, 13*(6), 543-558. doi: 10.1080/1043859032000124693

Koplow, D. N. (2006). *Biofuels--at what Cost?: Government Support for Ethanol and Biodiesel in the United States*: International Institute for Sustainable Development.

Kouvaritakis, N., Soria, A., & Isoard, S. (2000). Modelling energy technology dynamics: methodology for adaptive expectations models with learning by doing and learning by searching. *International Journal of Global Energy Issues, 14*(1-4), 104-115.

Langlois, R. N., & Steinmueller, W. E. (1999). The evolution of competitive advantage in the worldwide semiconductor industry, 1947-1996. *The Sources of Industrial Leadership*, 19-78.

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

Lasevoli, R. (2013). IPC vs. CPC compared. In T. I. Workshop (Ed.), *CPC Implementation Group*: World Intellectual Property Organization.

LexisNexis. (2015). SmartIndexing. *Academic - Wiki.* Retrieved 9/15/2015, 2015, from http://tinyurl.com/smrtidx

Liddy, E. D. (2001). Natural language processing *Encyclopedia of Library and Information Science* (2nd ed.). NY: Marcel Decker, Inc.

Lohwasser, R., & Madlener, R. (2013). Relating R&D and investment policies to CCS market diffusion through two-factor learning. *Energy Policy, 52*, 439-452. doi: 10.1016/j.enpol.2012.09.061

Loulou, R., Kanudia, A., Lehtila, A., & Goldstein, G. (2005). Documentation for the TIMES Model Part II.

Lundvall, B. A. (1992). National systems of innovation: An analytical framework. *London: Pinter*.

Lundvall, B. A. (2007). National innovation systems—analytical concept and development tool. *Industry and innovation, 14*(1), 95-119.

Lundvall, B. A., Johnson, B., Andersen, E. S., & Dalum, B. (2002). National systems of production, innovation and competence building. *Research Policy, 31*(2), 213-231. doi: 10.1016/s0048-7333(01)00137-8

Lynd, L. R., Laser, M. S., Bransby, D., Dale, B. E., Davison, B., Hamilton, R., . . . Wyman, C. E. (2008). How biotech can transform biofuels. *Nat Biotechnol, 26*(2), 169-172. doi: 10.1038/nbt0208-169

Malerba, F. (1985a). Demand structure and technological change: The case of the European semiconductor industry. *Research Policy, 14*(5), 283-297. doi: 10.1016/0048-7333(85)90010-1

Malerba, F. (1985b). *The semiconductor business: The economics of rapid growth and decline*: Univ of Wisconsin Press.

Malerba, F. (2002). Sectoral systems of innovation and production. *Research Policy, 31*(2), 247-264. doi: 10.1016/s0048-7333(01)00139-1

Malerba, F. (2005). Sectoral systems of innovation: a framework for linking innovation to the knowledge base, structure and dynamics of sectors. *Economics of Innovation and New Technology, 14*(1-2), 63-82. doi: 10.1080/1043859042000228688

McMahon, T. (2015). *Oil Prices 1946-Present*. Retrieved from: http://inflationdata.com/Inflation/Inflation_Rate/Historical_Oil_Prices_Table.asp

Meade, N., & Islam, T. (2006). Modelling and forecasting the diffusion of innovation – A 25-year review. *International Journal of Forecasting, 22*(3), 519-545. doi: 10.1016/j.ijforecast.2006.01.005

Melton, N., & Axsen, J. (2015). *Hype and Disappointment Cycles for Alternative Fuel Vehicle Technology: Analysis of US Media from 1980-2013.* Paper presented at the Transportation Research Board 94th Annual Meeting.

Mogee, M. E. (1991). Using patent data for technology analysis and planning. *Research Technology Management, 34*(4), 43.

Moser, P. (2013). Patents and innovation: evidence from economic history. *The Journal of Economic Perspectives*, 23-44.

Murphy, C. W., & Kendall, A. (2015). Life cycle analysis of biochemical cellulosic ethanol under multiple scenarios. *GCB Bioenergy, 7*(5), 1019-1033. doi: 10.1111/gcbb.12204

National Agricultural Statistical Service. (2015). *US Average Farm Price Received Database*. Retrieved from: http://farmdoc.illinois.edu/manage/uspricehistory/us_price_history.html

Negro, S. O., Hekkert, M. P., & Smits, R. E. (2007). Explaining the failure of the Dutch innovation system for biomass digestion—A functional analysis. *Energy Policy, 35*(2), 925-938. doi: 10.1016/j.enpol.2006.01.027

Nelson, R. (1993). National innovation systems: a comparative analysis. *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship*.

Nelson, R. R., & Winter, S. G. (1982). The Schumpeterian tradeoff revisited. *The American Economic Review*, 114-132.

Nordhaus, T., Shellenberger, M., & Navin, J. (2008). Fast, Clean, & Cheap: Cutting Global Warming's Gordian Knot. *Harvard Law and Policy Review*.

Nordhaus, W. (2009). doi: 10.3386/w14638

Peeters, C., & Pottelsberghe de la Potterie, B. (2006). Innovation strategy and the patenting behavior of firms. *Journal of Evolutionary Economics, 16*(1-2), 109-135. doi: 10.1007/s00191-005-0010-4

Pennock, D. M., Lawrence, S., Giles, C. L., & Nielsen, F. A. The real power of artificial markets.

Popp, D. (2001). Induced innovation and energy prices: National Bureau of Economic Research.

Popp, D. (2004). ENTICE: endogenous technological change in the DICE model of global warming. *Journal of Environmental Economics and Management, 48*(1), 742-768.

Popp, D. (2005). Lessons from patents: Using patents to measure technological change in environmental models. *Ecological Economics, 54*(2-3), 209-226. doi: 10.1016/j.ecolecon.2005.01.001

Popp, D., Juhl, T., & Johnson, D. K. (2003). Time in purgatory: Determinants of the grant lag for US patent applications: National Bureau of Economic Research.

Powers, C. (2014). Supporting the Plug-In Electric Vehicle Market Best Practices from State Programs. *Georgetown Climate Center, www. georgetownclimate. org/sites/GCC-Supporting-PEV-Market-December-2014. pdf.*

Qian, Y. (2007). Do National Patent Laws Stimulate Domestic Innovation in a Global Patenting Environment? A Cross-Country Analysis of Pharmaceutical Patent Protection, 1978–2002. *Review of Economics and Statistics, 89*(3), 436-453. doi: 10.1162/rest.89.3.436

Ray, R. (2006). Prediction Markets and the Financial "Wisdom of Crowds". *Journal of Behavioral Finance, 7*(1), 2-4. doi: 10.1207/s15427579jpfm0701_1

Romer, P. M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy, 94*(5), 1002-1037. doi: 10.2307/1833190

Romer, P. M. (1990). *Human capital and growth: theory and evidence.* Paper presented at the Carnegie-Rochester Conference Series on Public Policy.

Rubin, E. S., Yeh, S., Antes, M., Berkenpas, M., & Davison, J. (2007). Use of experience curves to estimate the future cost of power plants with CO2 capture. *International Journal of Greenhouse Gas Control, 1*(2), 188-197. doi: 10.1016/s1750-5836(07)00016-3

Sagar, A. D., & van der Zwaan, B. (2006). Technological innovation in the energy sector: R&D, deployment, and learning-by-doing. *Energy Policy, 34*(17), 2601-2608. doi: 10.1016/j.enpol.2005.04.012

Sakamoto, Y., Ishiguro, M., & Kitagawa, G. (1986). Akaike information criterion statistics. *Dordrecht, The Netherlands: D. Reidel*.

Sanderson, K. (2006). US biofuels: a field in ferment. *Nature, 444*(7120), 673-676. doi: 10.1038/444673a

Sarica, K., & Tyner, W. E. (2013). Analysis of US renewable fuels policies using a modified MARKAL model. *Renewable Energy, 50*, 701-709.

Savage, N. (2012). Gaining wisdom from crowds. *Communications of the ACM, 55*(3), 13. doi: 10.1145/2093548.2093553

Schmalensee, R. (2015). The Future of Solar Energy: A Personal Assessment. *Energy Economics*. doi: 10.1016/j.eneco.2015.08.012

Schmidt, C. W. (2007). ENVIRONMENT: California Out in Front. *Environmental Health Perspectives, 115*(3), A144-A147.

Schnepf, R., & Yacobucci, B. D. (2012). Renewable Fuel Standard (RFS): Overview and Issues Congressional Research Service.

Schubert, C. (2006). Can biofuels finally take center stage? *Nat Biotechnol, 24*(7), 777-784. doi: 10.1038/nbt0706-777

Schumpeter, J. A. (1934). *The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle* (Vol. 55): Transaction Publishers.

Schumpeter, J. A. (1942). *Socialism, capitalism and democracy*: Harper and Brothers.

Searchinger, T., Heimlich, R., Houghton, R. A., Dong, F., Elobeid, A., Fabiosa, J., . . . Yu, T.-H. (2008). Use of US croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science, 319*(5867), 1238-1240.

Sengers, F., Raven, R. P. J. M., & Van Venrooij, A. (2010). From riches to rags: Biofuels, media discourses, and resistance to sustainable energy technologies. *Energy Policy, 38*(9), 5013-5027. doi: http://dx.doi.org/10.1016/j.enpol.2010.04.030

Sims, R. E., Mabee, W., Saddler, J. N., & Taylor, M. (2010). An overview of second generation biofuel technologies. *Bioresour Technol, 101*(6), 1570-1580. doi: 10.1016/j.biortech.2009.11.046

Söderholm, P., & Klaassen, G. (2007). Wind power in Europe: a simultaneous innovation–diffusion model. *Environmental and resource economics, 36*(2), 163-190.

Solomon, B. D., Barnes, J. R., & Halvorsen, K. E. (2007). Grain and cellulosic ethanol: History, economics, and energy policy. *Biomass and Bioenergy, 31*(6), 416-425. doi: 10.1016/j.biombioe.2007.01.023

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economics, 70*(1), 65-94. doi: 10.2307/1884513

Solow, R. M. (1957). Technical Change and the Aggregate Production Function. *The Review of Economics and Statistics, 39*(3), 312. doi: 10.2307/1926047

Solow, R. M. (1994). Perspectives on Growth Theory. *The Journal of Economic Perspectives, 8*(1), 45-54. doi: 10.2307/2138150

Surowiecki, J. (2005). *The wisdom of crowds*: Anchor.

Suurs, R. A. A., & Hekkert, M. P. (2009). Cumulative causation in the formation of a technological innovation system: The case of biofuels in the Netherlands. *Technological Forecasting and Social Change, 76*(8), 1003-1020. doi: 10.1016/j.techfore.2009.03.002

Tichenor, P. J., Donohue, G. A., & Olien, C. N. (1970). Mass Media Flow and Differential Growth in Knowledge. *Public Opinion Quarterly, 34*(2), 159. doi: 10.1086/267786

Tigabu, A. D., Berkhout, F., & van Beukering, P. (2015). Technology innovation systems and technology diffusion: Adoption of bio-digestion in an emerging innovation system in Rwanda. *Technological Forecasting and Social Change, 90*, 318-330. doi: 10.1016/j.techfore.2013.10.011

Tyner, W. E. (2008). The US Ethanol and Biofuels Boom: Its Origins, Current Status, and Future Prospects. *BioScience, 58*(7), 646. doi: 10.1641/b580718

Tyner, W. E. (2012). Biofuels and agriculture: a past perspective and uncertain future. *International Journal of Sustainable Development & World Ecology, 19*(5), 389-394. doi: 10.1080/13504509.2012.691432

U.S. Department of Energy. (2012). Alternative Fuels Data Center - Federal and State Laws and Incentives. Retrieved 1/10/2012, from http://www.afdc.energy.gov/laws/

U.S. Energy Information Administration. (2012). NEMS Model Documentation 2012: Transportation Sector Module (pp. Maximum light-duty vehicle market penetration parameters (percent)).

U.S. Energy Information Administration. (2014). Annual Refinery Report.

U.S. Energy Information Administration. (2015a). Annual Energy Outlook 2015.

U.S. Energy Information Administration. (2015b). Monthly Biodiesel Production Report.

U.S. Energy Information Administration. (2015c). *State Energy Data System (SEDS): 1960-2013 (Complete)*. Retrieved from: http://www.eia.gov/state/seds/seds-data-complete.cfm

U.S. Environmental Protection Agency. (2004). *State Actions Banning MTBE (Statewide)*.

U.S. Patent and Trademark Office. (2013). FY 2013 Performance and Accountability Report

United States Congress. (1979). Gasohol: A Technical Memorandum (O. o. T. Assessment, Trans.).

Van de Ven, A. H., & Poole, M. S. (1990). Methods for studying innovation development in the Minnesota Innovation Research Program. *Organization science, 1*(3), 313-335.

Van den Bulte, C. (2000). New Product Diffusion Acceleration: Measurement and Analysis. *Marketing Science, 19*(4), 366-380. doi: 10.1287/mksc.19.4.366.11795

van der Zwaan, B. C. C., Gerlagh, R., G, Klaassen, & Schrattenholzer, L. (2002). Endogenous technological change in climate change modelling. *Energy Economics, 24*(1), 1-19. doi: 10.1016/s0140-9883(01)00073-1

van Sark, W. G. J. H. M., Alsema, E. A., Junginger, H. M., de Moor, H. H. C., & Schaeffer, G. J. (2008). Accuracy of progress ratios determined from experience curves: the case of crystalline silicon photovoltaic module technology development. *Progress in Photovoltaics: Research and Applications, 16*(5), 441-453. doi: 10.1002/pip.806

Ver Hoef, J. M., & Boveng, P. L. (2007). Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data? *Ecology, 88*(11), 2766-2772.

Wand, R., & Leuthold, F. (2011). Feed-in tariffs for photovoltaics: Learning by doing in Germany? *Applied Energy, 88*(12), 4387-4399. doi: 10.1016/j.apenergy.2011.05.015

Wareham, J., Levy, A., & Shi, W. (2004). Wireless diffusion and mobile computing: implications for the digital divide. *Telecommunications Policy, 28*(5-6), 439-457. doi: 10.1016/j.telpol.2003.11.005

Washington Department of Revenue. (2015). *Car Dealers and Leasing Companies Local Sales & Use Tax Rate Changes*. Washington Department of Revenue: Retrieved from http://dor.wa.gov/Docs/forms/ExcsTx/LocSalUseTx/MVETLocalSlsUseFlyer_15_Q4_.pdf.

Winebrake, J. J., & Farrell, A. (1997). The AFV credit program and its role in future AFV market development. *Transportation Research Part D: Transport and Environment, 2*(2), 125-132. doi: 10.1016/s1361-9209(97)00004-7

Witcover, J., Kessler, J., Eggert, A., & Yeh, S. (2015). The Low Carbon Fuel Standard. *Achieving California's Greenhouse Gas Goals: A Focus on Transportation*, 71.

Wright, T. P. (1936). Factors affecting the cost of airplanes. *Journal of the Aeronautical Sciences (Institute of the Aeronautical Sciences), 3*(4).

Wu, C.-H., Ken, Y., & Huang, T. (2010). Patent classification system using a new hybrid genetic algorithm support vector machine. *Applied Soft Computing, 10*(4), 1164-1177. doi: 10.1016/j.asoc.2009.11.033

Wu, J., & Shanley, M. T. (2009). Knowledge stock, exploration, and innovation: Research on the United States electromedical device industry. *Journal of business research, 62*(4), 474-483.

Yeh, S., Farrell, A., Plevin, R., Sanstad, A., & Weyant, J. (2008). Optimizing US mitigation strategies for the light-duty transportation sector: what we learn from a bottom-up model. *Environmental science & technology, 42*(22), 8202-8210.

Yeh, S., & Rubin, E. S. (2012). A review of uncertainties in technology experience curves. *Energy Economics, 34*(3), 762-771. doi: 10.1016/j.eneco.2011.11.006

Yeh, S., & Sperling, D. (2010). Low carbon fuel standards: Implementation scenarios and challenges. *Energy Policy, 38*(11), 6955-6965. doi: 10.1016/j.enpol.2010.07.012

Yeh, S., & Witcover, J. (2014a). Status Review of California's Low Carbon Fuel Standard - January 2014 Issue: University of California, Davis.

Yeh, S., & Witcover, J. (2014b). Status Review of California's Low Carbon Fuel Standard - July 2014 Issue: Institute of Transportation Studies.

Yeh, S., Witcover, J., & Bushnell, J. (2015). Status Review of California's Low Carbon Fuel Standard - April 2015 Issue (Revised Version): Institute of Transportation Studies.

Yeh, S., Witcover, J., & Kessler, J. (2013). Status Review of California's Low Carbon Fuel Standard - Spring 2013 (Revised Version): UC Davis Institute of Transportation Studies.

Yue, D., You, F., & Snyder, S. W. (2014). Biomass-to-bioenergy and biofuel supply chain optimization: Overview, key issues and challenges. *Computers & Chemical Engineering, 66*, 36-56. doi: 10.1016/j.compchemeng.2013.11.016

Zhuang, X., Engel, B. A., Lozano-Garcia, D. F., FernÁNdez, R. N., & Johannsen, C. J. (1994). Optimization of training data required for neuro-classification. *International Journal of Remote Sensing, 15*(16), 3271-3277. doi: 10.1080/01431169408954326

Zilberman, D., Hochman, G., Rajagopal, D., Sexton, S., & Timilsina, G. (2012). The Impact of Biofuels on Commodity Food Prices: Assessment of Findings. *American Journal of Agricultural Economics, 95*(2), 275-281. doi: 10.1093/ajae/aas037

Zucker, L. G., Darby, M. R., Furner, J., Liu, R. C., & Ma, H. (2007). Minerva unbound: Knowledge stocks, knowledge flows and new knowledge production. *Research Policy, 36*(6), 850-863.

Appendices

A. Supplemental Patent Information

Table A.1

Biopat Keyword List

Technology	**Keyword**	**Generation**	**Pathway**
Biodiesel	Fame	1	oil
Biodiesel	fatty acid methyl esters	1	oil
Biodiesel	fatty acid ethyl esters	1	oil
Biodiesel	free fatty acid	1	oil
Biodiesel	lipids as feedstock	1	oil
Biodiesel	lipids microbial organisms	2	algae
Biodiesel	fatty acyl-acp thioesterase	2	algae
Biodiesel	fatty acyl-coa/aldehyde reductase	2	algae
Biodiesel	fatty aldehyde decarbonylase	2	cyanobacteria
Biodiesel	acyl carrier protein	2	algae
Biodiesel	volatile fatty acids	2	waste digestion
Biodiesel	microbial lipids	2	algae
Biodiesel	microbial hosts	2	algae
Biodiesel	trichosporon	2	Fungi
Ethanol	agricultural feedstocks	1	fermentation

Ethanol	starch	1	fermentation
Ethanol	corn cobs	2	cellulosic
Ethanol	corn stover	2	cellulosic
Ethanol	cereal straw	2	cellulosic
Ethanol	forest harvest residues	2	cellulosic
Ethanol	husks	2	cellulosic
Biodiesel	chlorella vulgaris	2	algae
Biodiesel	spirulina maxima	2	algae
Biodiesel	nannochloropsis sp.	2	algae
Biodiesel	scenedesmus obliquus	2	algae
Biodiesel	dunaliella tertiolecta	2	algae
Biodiesel	scenedesmus dimorphus	2	algae
Biodiesel	eicosapentaenoic acid scenedesmus	2	algae
Ethanol	corn	1	fermentation
Ethanol	maize	1	fermentation
Ethanol	cassava	1	fermentation
Ethanol	grain	1	fermentation
Biodiesel	soybean	1	oil
Biodiesel	genetically engineered microbes	2	algae
Ethanol	genetically modified crops	1	fermentation
Ethanol	ligno-cellulosic	2	cellulosic
Ethanol	perennial grasses	2	cellulosic

Ethanol	forest	2	cellulosic
Ethanol	panicum virgatum 1	2	cellulosic
Ethanol	perennial plant	2	cellulosic
Ethanol	phalaris	2	cellulosic
Ethanol	alfalfa	2	cellulosic
Ethanol	reed canarygrass	2	cellulosic
Ethanol	fibrous plant materials	2	cellulosic
Ethanol	switchgrass	2	cellulosic
Ethanol	bark	2	cellulosic
Ethanol	wood shavings	2	cellulosic
Ethanol	chip boards	2	cellulosic
Ethanol	garden mulch	2	cellulosic
Ethanol	vegetative grasses	2	cellulosic
Ethanol	miscanthus	2	cellulosic
Ethanol	prairie grass	2	cellulosic
Ethanol	short rotation forest species	2	cellulosic
Ethanol	eucalyptus	2	cellulosic
Biodiesel	peanut	1	oil
Biodiesel	oil-bearing organisms	1	oil
Biodiesel	jatropha curcas	1	oil
Biodiesel	jatropha	1	oil
Biodiesel	babassu coconut	1	oil

Ethanol	helianthus tuberosus	1	fermentation
Biodiesel	oleaginous microorganisms	2	algae
Biodiesel	rhodotorula glutinis	2	yeast
Ethanol	medicago sativa l.	1	cellulosic
Ethanol	nut shells	1	cellulosic
Ethanol	sugarcane	1	fermentation
Ethanol	beet	1	fermentation
Ethanol	sorghum	1	fermentation
Ethanol	sugar esters	1	fermentation
Ethanol	bagasse	2	cellulosic
Ethanol	fermentable sugars	1	fermentation
Biodiesel	cooking oil	1	oil
Biodiesel	wet organic wastes	2	thermal chemical
Biodiesel	monosodium glutamate wastewater	2	algae
Ethanol	urban wood residues	2	cellulosic
Ethanol	ammonium	2	cellulosic
Biodiesel	animal waste	1	oil
Biodiesel	anlage	2	sludge
Biodiesel	excreta	2	sludge
Ethanol	feed mixture	1	fermentation
Ethanol	fibrobacter succinogenes	2	cellulosic
Biodiesel	kalium	1	oil

Biodiesel	chlorella emersonii	2	algae
Biodiesel	chlorella protothecoides	2	algae
Biodiesel	chlorella minutissima	2	algae
Biodiesel	dunaliella bioculata	2	algae
Biodiesel	dunaliella salina	2	algae
Biodiesel	microalgae oil	2	algae
Biodiesel	phaeodactylum tricornutum	2	algae
Biodiesel	vegetable oil	1	oil
Biodiesel	soya oil	1	oil
Biodiesel	untreated raw oils	1	oil
Biodiesel	oilseed rape	1	oil
Biodiesel	coconut oil	1	oil
Biodiesel	jojoba	1	oil
Biodiesel	canola oil	1	oil
Biodiesel	methanogenic bacteria	2	algae
Ethanol	poplars	2	cellulosic
Ethanol	lignin	2	cellulosic
Ethanol	cellulose	2	cellulosic
Ethanol	hemicellulose	2	cellulosic
Ethanol	wood process residues	2	cellulosic
Ethanol	wheat chaff	2	cellulosic
Biodiesel	animal fat	1	oil

Biodiesel	edible tallow	1	oil
Biodiesel	animal manure	1	oil
Biodiesel	granular sludge	2	sludge
Biodiesel	porcine pancreatic lipase	1	oil
Biodiesel	rapeseed	1	oil
Biodiesel	palm oil	1	oil
Ethanol	organic material	1	fermentation
Bodiesel	animal slurries	2	sludge
Ethanol	lignocellulose	2	cellulosic
Ethanol	liquid manure	2	fermentation
Biodiesel	microorganisms	2	algae
Ethanol	ruminococcus albus	2	cellulosic
Biodiesel	sewage	2	sludge
Biodiesel	siloxane	1	oil
Biodiesel	sulphide	1	oil
Biodiesel	digested sludge	2	sludge
Ethanol	fibrous material	2	cellulosic
Ethanol	hydrolysate	2	cellulosic
Ethanol	mesophilic bacteria	2	cellulosic
Biodiesel	microbial consortia	2	algae
Biodiesel	sludge	2	sludge
Biodiesel	treated wastewater	2	algae

Table A.2

Supplemental Table containing the top 20 most common CPC classifiers that correspond to Kessler and Sperling (2015) biofuel classification approach using NLP.

1G Classes	2G Classes	Ethanol	Biodiesel
A01H5/10	Y02E50/16	A01H5/10	A01H5/10
Y10S47/01	Y02E50/17	Y02E50/17	C11B1/00
A01H5/00	C12N9/2437	Y02E50/16	A23D9/00
A01H1/00	C12Y302/01004	C11B1/00	A01H5/00
C11B1/00	C12P7/10	C12N9/2437	A23L1/2003
A23D9/00	C12P19/14	Y10S47/01	A01H1/00
Y02E50/17	C11D3/38645	C12P19/14	Y10S47/01
C12N15/82	C12Y302/01021	C12P7/06	C12N15/82
C12N9/2417	C12Y302/01091	A23D9/00	A01H5/12
A01H1/02	C13K1/02	C12P7/10	C11B1/10
C12P7/06	C12N9/2445	A01H5/00	A23K1/14
C07K14/415	C12P7/06	C12Y302/01004	A01H4/00
C12P19/14	C12N15/8246	A01H1/00	A23K1/146
A01H4/00	C12P19/02	C12N9/2417	C12N5/04
A01H5/08	C12Y302/01008	C07K14/415	A01H1/02
A01H5/12	D21C5/005	C12Y302/01021	A01H5/04
A23L1/2003	Y02E50/343	C12N15/8245	C12N15/8247
C12N5/04	D06M16/003	A01H1/02	C11B1/04
C12N15/8245	A23K1/1653	C13K1/02	C12N15/8274
C12N9/2428	C10G2300/1014	C11D3/38645	Y02E50/13

Table A.3.

Regression results for 1st generation biofuel patents

	Dependent variable:					
	X1G.Patents		CPC_1G		GI_Eth	
	glm: quasipoisson *link = log*	*negative* *binomial*	*glm: quasipoisson* *link = log*	*negative* *binomial*	*glm: quasipoisson* *link = log*	*negative* *binomial*
	(1)	(2)	(3)	(4)	(5)	(6)
Other.Sector.Patents	-0.00000	0.00001	0.00001	0.00000	-0.00000	0.00000
	(0.00000)	(0.00001)	(0.00000)	(0.00001)	(0.00000)	(0.00000)
AgPatents	0.001***	0.001**	0.0005*	0.001**	0.001**	0.0004*
	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0002)	(0.0002)
Biotech	0.00000	0.00003	-0.0001***	-0.0001***	0.0001**	0.0001***
	(0.00003)	(0.00003)	(0.00004)	(0.00003)	(0.00002)	(0.00002)
Corn_1	0.245**	0.307***	-0.105	-0.178	0.231***	0.191**
	(0.093)	(0.118)	(0.107)	(0.114)	(0.083)	(0.087)
Oil_1	0.001	-0.005	0.011**	0.014***	0.003	0.002
	(0.004)	(0.005)	(0.004)	(0.004)	(0.003)	(0.003)
GE_Rev	1.188***	0.709**	0.435	0.744**	-0.325	-0.255
	(0.394)	(0.338)	(0.500)	(0.379)	(0.248)	(0.258)
RFS	1.052***	1.291***	0.793***	0.820***	0.749***	0.765***
	(0.224)	(0.277)	(0.266)	(0.267)	(0.208)	(0.219)
Constant	-0.755	-1.128	1.076	1.387	1.710**	2.165***
	(0.767)	(0.904)	(0.795)	(0.864)	(0.664)	(0.687)
Observations	34	34	34	34	34	34
Log Likelihood		-151.384		-127.168		-167.731
theta		11.579*** (3.830)		14.983** (6.331)		18.174*** (5.305)
Akaike Inf. Crit.		318.769		270.336		351.461
Note:						*p<0.1; **p<0.05; ***p<0.01

Table A.3

Regression results for 2nd generation biofuel patents

	Dependent variable:			
	X2G.Patents		CPC_2G	
	glm: quasipoisson link = log	*negative binomial*	*glm: quasipoisson link = log*	*negative binomial*
	(1)	(2)	(3)	(4)
Other.Sector.Patents	0.00000	0.00000	0.00000	0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
AgPatents	0.0001	0.0001	0.001***	0.001***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Biotech	0.0001**	0.0001***	-0.0001***	-0.0001***
	(0.00003)	(0.00003)	(0.00003)	(0.00003)
Corn_1	0.081	0.050	0.083	0.050
	(0.103)	(0.097)	(0.091)	(0.095)
Oil_1	0.011**	0.010***	0.017***	0.018***
	(0.004)	(0.004)	(0.003)	(0.003)
GE_Rev	0.070	0.158	-0.181	-0.176
	(0.352)	(0.297)	(0.393)	(0.313)
RFS	0.816***	0.864***	0.641***	0.650***
	(0.256)	(0.233)	(0.230)	(0.225)
Constant	0.579	0.916	-0.784	-0.536
	(0.782)	(0.748)	(0.687)	(0.744)
Observations	34	34	34	34
Log Likelihood		-113.291		-127.693
theta		25.684** (11.692)		22.685** (9.827)
Akaike Inf. Crit.		242.583		271.386
Note:			*p<0.1; **p<0.05; ***p<0.01	

B. Regression Result Summary for Biofuel TIS data

Table B.1

Policy Regression Results for Ethanol-related TIS data

	Dependent variable:						
	Legitimation	Entrepreneurial.Experimentation	Resource.Mobilization	Knowledge.development.and.diffusion	Development.of.Positive.External.Economies	Market.Formation	Influence.the.Direction.of.Search
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AFV.Manufacturer.Retrofitter	0.300***	0.310***	0.243**	0.323**		0.047	0.279**
	(0.084)	(0.099)	(0.106)	(0.146)		(0.139)	(0.130)
Grants	-0.173**				0.244**	0.195	-0.051
	(0.073)				(0.100)	(0.122)	(0.112)
Aftermarket.Conversions		-0.142*			0.308***	0.268**	0.083
		(0.081)			(0.108)	(0.131)	(0.119)
Alternative.Fuel.Producer		0.125*	0.136*			-0.035	0.089
		(0.068)	(0.077)			(0.100)	(0.092)
Climate.Change...Energy.Initiatives		-0.216*				-0.304	-0.284
		(0.119)				(0.207)	(0.192)
Exemptions						-0.086	-0.162
						(0.125)	(0.115)
Fleet.Purcha		-0.153**	-0.198**	-0.329***	-0.246***	-0.077	0.032

ser.Manager							
		(0.064)	(0.090)	(0.108)	(0.082)	(0.116)	(0.108)
Fuel.Economy...Efficiency		0.317***			0.289***	0.119	0.058
		(0.078)			(0.103)	(0.118)	(0.110)
Fuel.Production...Quality		-0.199***		-0.242**		0.062	0.077
		(0.072)		(0.100)		(0.108)	(0.100)
Fuel.Taxes						0.148	0.104
						(0.110)	(0.102)
Idle.Reduction	-0.156*	-0.172		-0.330**	-0.305**	-0.383**	-0.181
	(0.087)	(0.105)		(0.146)	(0.134)	(0.152)	(0.141)
Loans.and.Leases	-0.160				-0.339**	-0.289	-0.235
	(0.106)				(0.153)	(0.178)	(0.162)
Rebates						0.293	0.070
						(0.235)	(0.211)
Registration...Licensing	-0.130**					-0.124	-0.146
	(0.061)					(0.113)	(0.104)
Tax.Incentives		0.214***				0.163	0.131
		(0.067)				(0.101)	(0.094)
Fueling...TSE.Infrastructure.Owner			-0.146*			0.215*	-0.074

			(0.080)			(0.119)	(0.109)
Vehicle.Owner.Driver				-0.299***		-0.168	0.023
				(0.098)		(0.109)	(0.103)
Renewable.Fuel.Standard...Mandate					0.307***	0.458***	0.175
					(0.119)	(0.134)	(0.123)
stateAlaska	1.950***	0.860***	1.480***	0.358	1.123***	1.986***	2.565***
	(0.153)	(0.196)	(0.205)	(0.287)	(0.239)	(0.267)	(0.243)
stateArizona	1.291***	0.812***	0.516**	0.764***	1.297***	1.278***	0.931***
	(0.152)	(0.183)	(0.206)	(0.276)	(0.221)	(0.239)	(0.226)
stateArkansas	0.557***	0.128	0.234	0.155	0.125	0.431	0.702***
	(0.165)	(0.206)	(0.224)	(0.333)	(0.246)	(0.277)	(0.253)
stateCalifornia	3.641***	3.239***	2.815***	3.791***	3.279***	3.530***	3.551***
	(0.151)	(0.169)	(0.189)	(0.254)	(0.215)	(0.257)	(0.238)
stateColorado	1.511***	1.600***	0.957***	2.277***	1.182***	1.387***	1.273***
	(0.155)	(0.192)	(0.202)	(0.296)	(0.225)	(0.277)	(0.259)
stateConnecticut	1.146***	0.362*	-0.190	0.598**	-0.073	1.161***	0.828***
	(0.156)	(0.192)	(0.219)	(0.286)	(0.252)	(0.257)	(0.245)
stateDelaware	0.165	-0.127	-0.104	0.071	0.250	0.289	0.370
	(0.158)	(0.196)	(0.218)	(0.282)	(0.237)	(0.278)	(0.259)

stateDistrict of Columbia	2.636***	1.389***	2.292***	2.649***	2.146***	2.207***	2.629***
	(0.158)	(0.180)	(0.198)	(0.263)	(0.224)	(0.268)	(0.246)
stateFlorida	1.720***	1.389***	1.388***	1.699***	1.427***	1.062***	1.681***
	(0.152)	(0.179)	(0.198)	(0.267)	(0.221)	(0.246)	(0.223)
stateGeorgia	0.879***	0.810***	0.349*	1.291***	0.437*	0.151	0.495*
	(0.154)	(0.195)	(0.209)	(0.276)	(0.249)	(0.300)	(0.277)
stateHawaii	0.159	0.130	0.161	0.204	-0.369	0.294	0.388
	(0.170)	(0.213)	(0.233)	(0.325)	(0.266)	(0.309)	(0.290)
stateIdaho	0.599***	0.593***	0.103	0.635**	-0.362	0.226	0.610***
	(0.155)	(0.188)	(0.214)	(0.282)	(0.252)	(0.258)	(0.233)
stateIllinois	2.668***	2.523***	2.326***	2.447***	1.933***	2.650***	2.774***
	(0.149)	(0.173)	(0.194)	(0.256)	(0.222)	(0.239)	(0.221)
stateIndiana	1.522***	1.261***	1.024***	1.233***	1.062***	1.450***	1.315***
	(0.155)	(0.188)	(0.204)	(0.293)	(0.230)	(0.249)	(0.234)
stateIowa	2.967***	2.500***	2.501***	2.198***	2.362***	2.943***	3.155***
	(0.149)	(0.168)	(0.187)	(0.258)	(0.212)	(0.234)	(0.216)
stateKansas	1.940***	2.050***	1.353***	1.563***	1.308***	1.343***	1.875***
	(0.152)	(0.183)	(0.196)	(0.278)	(0.235)	(0.324)	(0.290)
stateKentucky	0.608***	-0.261	0.457**	0.005	0.888***	0.920***	0.848***
	(0.16	(0.214)	(0.232)	(0.316)	(0.230)	(0.274	(0.259)

	8)					)	
stateLouisiana	1.214***	1.090***	0.754***	1.134***	0.962***	1.232***	0.976***
	(0.153)	(0.179)	(0.204)	(0.278)	(0.238)	(0.254)	(0.240)
stateMaine	0.527***	-0.755***	-0.595***	-0.141	0.233	0.499*	-0.162
	(0.155)	(0.209)	(0.230)	(0.316)	(0.237)	(0.268)	(0.263)
stateMaryland	1.082***	0.535***	0.245	0.866***	0.888***	1.046***	0.958***
	(0.155)	(0.186)	(0.211)	(0.282)	(0.232)	(0.249)	(0.233)
stateMassachusetts	1.744***	1.086***	1.113***	2.067***	1.414***	1.391***	1.772***
	(0.150)	(0.177)	(0.198)	(0.251)	(0.220)	(0.239)	(0.219)
stateMichigan	1.662***	1.424***	1.033***	1.409***	1.788***	1.284***	1.239***
	(0.152)	(0.176)	(0.200)	(0.257)	(0.216)	(0.245)	(0.229)
stateMinnesota	2.511***	2.509***	2.062***	2.493***	2.163***	2.205***	2.398***
	(0.151)	(0.173)	(0.193)	(0.271)	(0.223)	(0.272)	(0.250)
stateMississippi	0.655***	1.039***	0.460**	0.931***	0.062	0.288	0.630***
	(0.155)	(0.181)	(0.208)	(0.275)	(0.245)	(0.259)	(0.236)
stateMissouri	1.886***	1.626***	1.687***	1.792***	1.581***	1.920***	2.127***
	(0.157)	(0.177)	(0.200)	(0.271)	(0.229)	(0.262)	(0.242)
stateMontana	0.745***	0.352*	0.304	0.583**	-0.270	0.040	1.188***
	(0.154)	(0.205)	(0.209)	(0.282)	(0.274)	(0.288)	(0.252)
stateNebraska	2.260***	2.530***	1.980***	0.883***	1.346***	2.068***	2.569***

	(0.152)	(0.171)	(0.195)	(0.274)	(0.229)	(0.246)	(0.226)
stateNevada	0.579***	0.455**	0.174	0.354	1.522***	1.586***	0.253
	(0.155)	(0.186)	(0.211)	(0.307)	(0.218)	(0.250)	(0.250)
stateNew Hampshire	1.145***	-0.174	-0.579**	0.097	-0.049	0.912***	1.130***
	(0.154)	(0.201)	(0.234)	(0.316)	(0.251)	(0.253)	(0.231)
stateNew Jersey	1.111***	0.720***	0.284	0.692**	0.549**	0.903***	0.658***
	(0.152)	(0.182)	(0.209)	(0.280)	(0.231)	(0.244)	(0.230)
stateNew Mexico	1.449***	1.151***	0.725***	1.895***	0.814***	1.124***	1.327***
	(0.165)	(0.205)	(0.204)	(0.289)	(0.261)	(0.323)	(0.298)
stateNew York	2.915***	2.611***	1.860***	2.820***	2.424***	2.902***	2.570***
	(0.160)	(0.189)	(0.211)	(0.278)	(0.234)	(0.265)	(0.248)
stateNorth Carolina	1.001***	1.184***	0.805***	1.371***	0.596**	0.151	0.827***
	(0.155)	(0.192)	(0.211)	(0.281)	(0.245)	(0.295)	(0.265)
stateNorth Dakota	1.583***	1.749***	1.321***	1.143***	0.816***	1.040***	1.783***
	(0.152)	(0.174)	(0.196)	(0.271)	(0.228)	(0.247)	(0.223)
stateOhio	1.335***	1.478***	1.187***	1.651***	1.265***	0.851**	1.063***
	(0.154)	(0.197)	(0.242)	(0.274)	(0.288)	(0.355)	(0.329)
stateOklahoma	0.780***	1.001***	0.500**	0.888***	0.679***	0.514*	0.779***
	(0.174)	(0.209)	(0.227)	(0.309)	(0.247)	(0.297)	(0.275)
stateOregon	0.52	0.389**	0.272	0.667**	0.453*	0.290	0.352

	9***						
	(0.160)	(0.194)	(0.231)	(0.296)	(0.256)	(0.281)	(0.261)
statePennsylvania	1.459***	1.183***	1.147***	1.464***	0.977***	1.071***	1.243***
	(0.157)	(0.177)	(0.198)	(0.275)	(0.228)	(0.285)	(0.264)
stateRhode Island	-0.063	-1.329***	-1.038***	-1.327***	-1.267***	0.006	0.051
	(0.161)	(0.224)	(0.248)	(0.371)	(0.288)	(0.269)	(0.249)
stateSouth Carolina	0.474***	0.371**	0.613***	0.880***	0.065	-0.254	0.006
	(0.156)	(0.188)	(0.204)	(0.276)	(0.242)	(0.268)	(0.243)
stateSouth Dakota	1.725***	1.930***	1.414***	0.863***	0.860***	1.538***	1.910***
	(0.151)	(0.177)	(0.195)	(0.276)	(0.226)	(0.249)	(0.229)
stateTennessee	0.941***	0.756***	0.623***	1.626***	0.785***	0.908***	0.524**
	(0.154)	(0.184)	(0.208)	(0.270)	(0.231)	(0.267)	(0.258)
stateTexas	2.772***	2.772***	2.032***	2.681***	2.079***	2.528***	2.760***
	(0.149)	(0.180)	(0.192)	(0.260)	(0.230)	(0.259)	(0.240)
stateUtah	0.076	0.047	-0.406*	-0.173	0.058	-0.213	-0.019
	(0.164)	(0.202)	(0.231)	(0.319)	(0.254)	(0.286)	(0.258)
stateVermont	-0.295*	-1.543***	-1.536***	-1.638***	-0.443	-0.203	-0.548*
	(0.172)	(0.251)	(0.277)	(0.422)	(0.272)	(0.308)	(0.297)
stateVirginia	1.496***	1.108***	0.551***	1.483***	0.730***	1.045***	1.107***

	(0.178)	(0.189)	(0.205)	(0.266)	(0.249)	(0.317)	(0.297)
stateWashington	3.742***	2.555***	3.416***	3.299***	3.095***	3.724***	4.047***
	(0.147)	(0.172)	(0.188)	(0.248)	(0.217)	(0.237)	(0.219)
stateWest Virginia	-0.358**	-1.011***	-0.770***	0.604*	-0.111	-0.510	-1.015***
	(0.163)	(0.240)	(0.252)	(0.347)	(0.254)	(0.324)	(0.308)
stateWisconsin	1.662***	1.392***	1.203***	1.612***	1.082***	1.413***	1.695***
	(0.157)	(0.176)	(0.199)	(0.257)	(0.225)	(0.256)	(0.236)
stateWyoming	0.389**	0.203	-0.013	-0.170	-0.146	0.035	0.506*
	(0.162)	(0.188)	(0.215)	(0.306)	(0.246)	(0.288)	(0.258)
as.factor(year)1997	0.329***	-0.109	0.275	0.894***	0.918***	0.921***	0.610***
	(0.094)	(0.134)	(0.173)	(0.238)	(0.170)	(0.187)	(0.155)
as.factor(year)1998	0.352***	-0.275**	0.139	0.017	1.064***	1.286***	0.243
	(0.094)	(0.138)	(0.176)	(0.274)	(0.168)	(0.182)	(0.162)
as.factor(year)1999	1.118***	0.241*	0.354**	0.893***	1.453***	1.773***	0.794***
	(0.090)	(0.127)	(0.170)	(0.238)	(0.163)	(0.176)	(0.153)
as.factor(year)2000	1.590***	0.730***	1.268***	0.503**	1.811***	2.353***	1.601***
	(0.088)	(0.120)	(0.154)	(0.251)	(0.159)	(0.172)	(0.145)
as.factor(year)2001	1.491***	1.062***	1.363***	0.744***	1.653***	2.458***	1.400***
	(0.089)	(0.117)	(0.153)	(0.242)	(0.161)	(0.173)	(0.148)

as.factor(year)2002	1.578***	0.883***	1.517***	1.034***	0.691***	1.963***	1.986***
	(0.089)	(0.119)	(0.151)	(0.234)	(0.174)	(0.177)	(0.144)
as.factor(year)2003	2.134***	1.429***	1.857***	1.065***	1.181***	2.665***	2.616***
	(0.088)	(0.115)	(0.149)	(0.235)	(0.167)	(0.174)	(0.143)
as.factor(year)2004	2.264***	1.559***	2.121***	2.195***	1.697***	2.531***	2.522***
	(0.088)	(0.114)	(0.147)	(0.216)	(0.162)	(0.175)	(0.143)
as.factor(year)2005	2.903***	2.485***	3.062***	2.795***	2.830***	3.590***	3.175***
	(0.088)	(0.110)	(0.143)	(0.213)	(0.155)	(0.173)	(0.143)
as.factor(year)2006	3.775***	3.766***	3.787***	3.982***	3.895***	4.175***	3.909***
	(0.089)	(0.109)	(0.143)	(0.210)	(0.154)	(0.174)	(0.145)
as.factor(year)2007	4.059***	3.859***	4.311***	4.613***	4.361***	4.750***	4.277***
	(0.091)	(0.111)	(0.144)	(0.213)	(0.155)	(0.180)	(0.151)
as.factor(year)2008	4.080***	3.955***	4.227***	4.650***	4.312***	4.495***	4.088***
	(0.094)	(0.115)	(0.147)	(0.217)	(0.158)	(0.188)	(0.158)
as.factor(year)2009	3.444***	3.426***	3.712***	4.268***	3.604***	3.732***	3.366***
	(0.097)	(0.120)	(0.153)	(0.222)	(0.162)	(0.195)	(0.166)
as.factor(year)2010	3.482***	3.418***	3.711***	4.217***	3.396***	3.295***	3.534***
	(0.100)	(0.124)	(0.158)	(0.226)	(0.166)	(0.205)	(0.174)
Acquisition... Fuel.Use	0.078		0.226***	0.241**		-0.064	-0.088
	(0.05		(0.085)	(0.104)		(0.109	(0.101)

	3)					)	
Air.Quality... Emissions		-0.213			-0.320**	0.131	0.029
		(0.131)			(0.158)	(0.214)	(0.196)
Alternative. Fuel.Dealer						-0.273**	-0.145
						(0.107)	(0.098)
Alternative. Fuel.Purchaser				0.224**		0.192*	-0.004
				(0.095)		(0.109)	(0.100)
Constant	1.021***	-0.107	-0.716***	-2.148***	-0.843***	-1.226***	-0.703***
	(0.133)	(0.166)	(0.199)	(0.292)	(0.221)	(0.242)	(0.212)
Observations	765	765	765	765	765	765	765
Log Likelihood	-3,811.988	-2,590.410	-2,364.494	-1,729.101	-2,407.157	-2,670.215	-2,779.524
theta	7.589***(0.471)	9.136*** (0.835)	7.127*** (0.677)	7.149*** (0.865)	5.382*** (0.449)	4.088*** (0.292)	4.796*** (0.364)
Akaike Inf. Crit.	7,765.976	5,328.820	4,870.987	3,602.202	4,960.313	5,514.429	5,733.048
Note:							*p<0.1; **p<0.05; ***p<0.01

C. Regression Result Summary for Electric Vehicles

Table C.1

Electric Vehicle Sales Regression Results

	Dependent variable:	
	Vehicles_Sold	
	negative binomial	*glm: quasipoisson link = log*
	(1)	(2)
Vehicles_1	0.0005***	0.0002***
	(0.0001)	(0.00002)
Res_Elec	0.054*	0.039***
	(0.028)	(0.008)
VMT_Rural	0.0004***	0.0002***
	(0.0001)	(0.00004)
StateCALIFORNIA	1.933***	2.663**
	(0.447)	(1.219)
StateCOLORADO	-0.187	-0.205
	(0.433)	(1.781)
StateCONNECTICUT	-0.072	-0.129
	(0.500)	(1.664)
StateFLORIDA	0.761**	0.999
	(0.378)	(1.361)
StateGEORGIA	-0.234	-0.061
	(0.426)	(1.662)
StateILLINOIS	0.465	0.588
	(0.381)	(1.461)
StateINDIANA	-0.112	-0.070
	(0.424)	(1.715)
StateMAINE	-2.847***	-2.879
	(1.067)	(4.918)
StateMARYLAND	0.820**	0.797
	(0.385)	(1.417)

StateMASSACHUSETTS	-0.115	-0.224
	(0.462)	(1.782)
StateMICHIGAN	1.286***	1.451
	(0.359)	(1.305)
StateMINNESOTA	-0.827*	-0.750
	(0.487)	(2.066)
StateNEW JERSEY	0.670	0.613
	(0.424)	(1.454)
StateNEW MEXICO	-1.063*	-1.115
	(0.558)	(2.444)
StateNEW YORK	0.108	0.276
	(0.437)	(1.478)
StateNORTH CAROLINA	0.315	0.438
	(0.393)	(1.505)
StateOREGON	-0.058	-0.117
	(0.436)	(1.781)
StateRHODE ISLAND	-1.991**	-2.120
	(0.808)	(3.579)
StateTENNESSEE	0.241	0.356
	(0.402)	(1.540)
StateTEXAS	0.859*	1.526
	(0.493)	(1.270)
StateVERMONT	-2.739**	-2.830
	(1.070)	(4.918)
StateVIRGINIA	0.627*	0.754
	(0.380)	(1.422)
StateWASHINGTON	0.831**	0.753
	(0.386)	(1.477)
as.factor(Year)2011	3.390***	3.415***
	(0.336)	(1.212)
as.factor(Year)2012	3.770***	3.790***
	(0.343)	(1.208)
as.factor(Year)2013	4.344***	4.396***
	(0.335)	(1.200)

as.factor(Year)2014	4.478***	4.559***
	(0.339)	(1.199)
StateCALIFORNIA:as.factor(Year)2011	-0.782*	-0.738
	(0.411)	(1.236)
StateCOLORADO:as.factor(Year)2011	-0.517	-0.543
	(0.492)	(1.822)
StateCONNECTICUT:as.factor(Year)2011	-1.572***	-1.608
	(0.481)	(1.743)
StateFLORIDA:as.factor(Year)2011	-0.822*	-0.847
	(0.427)	(1.390)
StateGEORGIA:as.factor(Year)2011	-0.461	-0.485
	(0.473)	(1.696)
StateILLINOIS:as.factor(Year)2011	-0.712	-0.732
	(0.442)	(1.493)
StateINDIANA:as.factor(Year)2011	-0.865*	-0.870
	(0.483)	(1.762)
StateMAINE:as.factor(Year)2011	-0.317	-0.353
	(1.100)	(5.031)
StateMARYLAND:as.factor(Year)2011	-1.318***	-1.330
	(0.437)	(1.459)
StateMASSACHUSETTS:as.factor(Year)2011	-0.954*	-0.967
	(0.496)	(1.836)
StateMICHIGAN:as.factor(Year)2011	-1.474***	-1.489
	(0.420)	(1.338)
StateMINNESOTA:as.factor(Year)2011	-0.042	-0.068
	(0.539)	(2.101)
StateNEW JERSEY:as.factor(Year)2011	-1.107**	-1.127
	(0.442)	(1.493)
StateNEW MEXICO:as.factor(Year)2011	-0.961	-0.976
	(0.615)	(2.533)
StateNEW YORK:as.factor(Year)2011	-0.856*	-0.884
	(0.445)	(1.513)
StateNORTH CAROLINA:as.factor(Year)2011	-0.708	-0.711
	(0.448)	(1.539)

StateOREGON:as.factor(Year)2011	0.346	0.376
	(0.490)	(1.805)
StateRHODE ISLAND:as.factor(Year)2011	-1.508*	-1.561
	(0.864)	(3.840)
StateTENNESSEE:as.factor(Year)2011	-0.716	-0.706
	(0.454)	(1.574)
StateTEXAS:as.factor(Year)2011	-1.184***	-1.196
	(0.414)	(1.291)
StateVERMONT:as.factor(Year)2011	-0.214	-0.223
	(1.096)	(5.014)
StateVIRGINIA:as.factor(Year)2011	-1.301***	-1.291
	(0.437)	(1.463)
StateWASHINGTON:as.factor(Year)2011	0.263	0.319
	(0.442)	(1.498)
StateCALIFORNIA:as.factor(Year)2012	-0.484	-0.186
	(0.447)	(1.232)
StateCOLORADO:as.factor(Year)2012	0.340	0.358
	(0.499)	(1.801)
StateCONNECTICUT:as.factor(Year)2012	-0.321	-0.350
	(0.487)	(1.690)
StateFLORIDA:as.factor(Year)2012	-0.086	-0.082
	(0.437)	(1.379)
StateGEORGIA:as.factor(Year)2012	0.122	0.145
	(0.481)	(1.682)
StateILLINOIS:as.factor(Year)2012	0.098	0.109
	(0.452)	(1.479)
StateINDIANA:as.factor(Year)2012	-0.732	-0.727
	(0.491)	(1.749)
StateMAINE:as.factor(Year)2012	1.516	1.485
	(1.086)	(4.936)
StateMARYLAND:as.factor(Year)2012	-0.279	-0.243
	(0.447)	(1.437)
StateMASSACHUSETTS:as.factor(Year)2012	0.519	0.523
	(0.499)	(1.800)

StateMICHIGAN:as.factor(Year)2012	-0.841*	-0.754
	(0.431)	(1.325)
StateMINNESOTA:as.factor(Year)2012	0.746	0.778
	(0.545)	(2.083)
StateNEW JERSEY:as.factor(Year)2012	-0.151	-0.143
	(0.451)	(1.474)
StateNEW MEXICO:as.factor(Year)2012	-0.490	-0.494
	(0.616)	(2.492)
StateNEW YORK:as.factor(Year)2012	0.442	0.488
	(0.455)	(1.494)
StateNORTH CAROLINA:as.factor(Year)2012	-0.339	-0.345
	(0.458)	(1.528)
StateOREGON:as.factor(Year)2012	0.454	0.496
	(0.499)	(1.800)
StateRHODE ISLAND:as.factor(Year)2012	0.171	0.127
	(0.828)	(3.623)
StateTENNESSEE:as.factor(Year)2012	-0.680	-0.678
	(0.464)	(1.568)
StateTEXAS:as.factor(Year)2012	-1.063**	-1.030
	(0.425)	(1.285)
StateVERMONT:as.factor(Year)2012	1.145	1.150
	(1.086)	(4.940)
StateVIRGINIA:as.factor(Year)2012	-0.855*	-0.807
	(0.445)	(1.447)
StateWASHINGTON:as.factor(Year)2012	0.192	0.250
	(0.453)	(1.495)
StateCALIFORNIA:as.factor(Year)2013	-1.150**	-0.340
	(0.560)	(1.225)
StateCOLORADO:as.factor(Year)2013	0.289	0.297
	(0.487)	(1.790)
StateCONNECTICUT:as.factor(Year)2013	-0.162	-0.203
	(0.474)	(1.673)
StateFLORIDA:as.factor(Year)2013	-0.242	-0.231
	(0.425)	(1.369)

StateGEORGIA:as.factor(Year)2013	1.089**	1.200
	(0.469)	(1.668)
StateILLINOIS:as.factor(Year)2013	0.030	0.007
	(0.441)	(1.470)
StateINDIANA:as.factor(Year)2013	-0.883*	-0.892
	(0.479)	(1.732)
StateMAINE:as.factor(Year)2013	1.322	1.255
	(1.081)	(4.928)
StateMARYLAND:as.factor(Year)2013	-0.562	-0.561
	(0.434)	(1.427)
StateMASSACHUSETTS:as.factor(Year)2013	0.404	0.413
	(0.488)	(1.790)
StateMICHIGAN:as.factor(Year)2013	-0.891**	-0.826
	(0.419)	(1.314)
StateMINNESOTA:as.factor(Year)2013	0.220	0.209
	(0.535)	(2.076)
StateNEW JERSEY:as.factor(Year)2013	-0.034	-0.017
	(0.440)	(1.462)
StateNEW MEXICO:as.factor(Year)2013	-0.743	-0.781
	(0.604)	(2.472)
StateNEW YORK:as.factor(Year)2013	0.411	0.488
	(0.442)	(1.484)
StateNORTH CAROLINA:as.factor(Year)2013	-0.721	-0.737
	(0.445)	(1.518)
StateOREGON:as.factor(Year)2013	0.609	0.645
	(0.487)	(1.789)
StateRHODE ISLAND:as.factor(Year)2013	0.089	0.022
	(0.817)	(3.600)
StateTENNESSEE:as.factor(Year)2013	-0.744*	-0.759
	(0.451)	(1.553)
StateTEXAS:as.factor(Year)2013	-1.305***	-1.237
	(0.412)	(1.275)
StateVERMONT:as.factor(Year)2013	1.512	1.507
	(1.079)	(4.925)

StateVIRGINIA:as.factor(Year)2013	-0.923**	-0.908
	(0.432)	(1.434)
StateWASHINGTON:as.factor(Year)2013	0.476	0.574
	(0.441)	(1.484)
StateCALIFORNIA:as.factor(Year)2014	-1.664**	-0.427
	(0.701)	(1.226)
StateCOLORADO:as.factor(Year)2014	0.064	0.056
	(0.492)	(1.791)
StateCONNECTICUT:as.factor(Year)2014	-0.725	-0.764
	(0.475)	(1.676)
StateFLORIDA:as.factor(Year)2014	-0.308	-0.293
	(0.430)	(1.369)
StateGEORGIA:as.factor(Year)2014	1.364***	1.607
	(0.480)	(1.667)
StateILLINOIS:as.factor(Year)2014	-0.433	-0.461
	(0.445)	(1.470)
StateINDIANA:as.factor(Year)2014	-0.817*	-0.832
	(0.483)	(1.730)
StateMAINE:as.factor(Year)2014	0.724	0.642
	(1.083)	(4.933)
StateMARYLAND:as.factor(Year)2014	-0.883**	-0.894
	(0.439)	(1.428)
StateMASSACHUSETTS:as.factor(Year)2014	0.035	0.027
	(0.494)	(1.791)
StateMICHIGAN:as.factor(Year)2014	-1.044**	-0.979
	(0.423)	(1.314)
StateMINNESOTA:as.factor(Year)2014	-0.318	-0.333
	(0.540)	(2.080)
StateNEW JERSEY:as.factor(Year)2014	-0.558	-0.598
	(0.445)	(1.464)
StateNEW MEXICO:as.factor(Year)2014	-0.748	-0.795
	(0.608)	(2.469)
StateNEW YORK:as.factor(Year)2014	0.212	0.309
	(0.446)	(1.484)

StateNORTH CAROLINA:as.factor(Year)2014	-0.986**	-0.998
	(0.450)	(1.518)
StateOREGON:as.factor(Year)2014	0.384	0.406
	(0.492)	(1.789)
StateRHODE ISLAND:as.factor(Year)2014	-0.329	-0.388
	(0.820)	(3.605)
StateTENNESSEE:as.factor(Year)2014	-1.068**	-1.111
	(0.456)	(1.555)
StateTEXAS:as.factor(Year)2014	-1.458***	-1.348
	(0.418)	(1.274)
StateVERMONT:as.factor(Year)2014	1.024	0.990
	(1.081)	(4.928)
StateVIRGINIA:as.factor(Year)2014	-1.145***	-1.134
	(0.438)	(1.434)
StateWASHINGTON:as.factor(Year)2014	0.290	0.367
	(0.446)	(1.484)
Constant	-0.683	-0.299
	(0.415)	(1.196)
Observations	1,368	1,368
Log Likelihood	-5,955.114	
theta	3.419*** (0.177)	
Akaike Inf. Crit.	12,156.230	
Note:	*p**p***p<0.01	