Necessity Is the Mother of Invention:

How Minimum Wages Affect Innovation

Matias Vidnes Seip

Master's thesis

Master's thesis to complete the degree

Master's in economics

University of Bergen, Department of Economics

June 2023

UNIVERSITETET I BERGEN

Preface

Through the writing of this thesis, I have learned a great deal – not just about the research topic, but about the process of doing research itself. This would not have been possible without the good help of others, for which I am immensely thankful.

First and foremost, I am profoundly grateful to my supervisor, Marc Goñi, whose competence and patience contributed considerably to my experience. His guidance has been instrumental in the successful completion of this master's thesis.

I would also like to thank my friend, Balder Blinkenberg, who kindly lent his time to me and whose insightful comments and constructive critiques were greatly beneficial in improving the quality of this work.

Lastly, I would like to extend my gratitude to my family and my friends for enriching this period of my life.

I am solely responsible for any interpretations or errors in the thesis.

Bergen, May 26th, 2023

Matias Vidnes Seip

Abstract

This master's thesis investigates the impact of minimum wages on innovation in the United States. I construct a panel dataset with minimum wages and patent production across US states between 1983 and 2015. My estimates are based on a difference-in-differences approach that exploits the different timing of minimum wage raises across US states and the fact that federal minimum wage is not inflation adjusted. Two-way fixed-effects' estimates suggest that a 1 USD increase in the minimum wage is associated with approximately a 9 percent increase in patent production. Accounting for heterogeneous treatment effects suggests that this effect is potentially larger. Altogether, my results provide support for the Habakkuk thesis that labor scarcity encourages technological innovation. It also provides insights into recent debates surrounding the causes of automation as well as calls for the minimum wage to be raised in the United States.

All calculations and estimations have been conducted using Stata version 17.0.

Table of Contents

Tables

Figures

1 Introduction

In recent years, the automation of human labor has been a hotly debated topic in the public discourse, as well as in the economics literature. The emergence of sophisticated technologies like artificial intelligence and robotics brings with it the promise of unprecedented productivity and efficiency in various sectors. However, this wave of automation also raises some difficult societal questions. Some argue that automation could potentially lead to job displacement and increased income disparity, while others view it as a driving force for new job creation and economic growth.

In the public debate, the effect of wage costs on technological innovations has perhaps not featured prominently. Nevertheless, economic theory suggests that automation and labor costs are closely knitted. The argument that labor scarcity can encourage capital use goes back decades and has been raised by prominent economic thinkers like John Hicks (1932) and John Habakkuk (1962). More recently Daron Acemoğlu has also made this connection in several papers (Acemoğlu & Finkelstein, 2008; Acemoğlu, 2010; Acemoğlu & Restrepo, 2020). Put simply, the argument goes that when labor is expensive, employers instead turn to technology as a substitute. Accepting this statement as true however, requires further scrutiny, and investigating its legitimacy will be the focal point of this thesis. The research question of the thesis is:

Can a wage floor lead to increased innovation?

To investigate the research question, I employ a difference-in-differences (DiD) approach using a panel dataset that I have constructed from minimum wage rates and the number of patents granted in the separate US states between 1983 and 2015. The technique compares patent production before and after a state raises its minimum wage, with states who did not raise their minimum wage serving as a control group. Several covariates are also included in the panel dataset to ensure that shocks hitting the control and treatment group differently were not the cause of the differences.

The DiD method used in this thesis has commonly been utilized to investigate the impact of minimum wages on employment levels and has been adapted in this study to explore its influence on innovation, represented by patent production. This reworking of a well-established method offers a novel angle of investigation into the relationship between wage floors and technological progress. As the effect of minimum wages on innovation is less known, this thesis can help shed light on this aspect of the policy. In this regard, the thesis is timely considering recent calls to raise the federal minimum wage in the United States (Figueroa, 2022).

Previous studies have investigated a relationship similar to the one studied in this thesis (see section 2.3). As such, the main way through which this thesis contributes to the existing body of knowledge is through its broad scope. Rather than focusing on specific industries or types of patents – the common approach of previous research – this study encompasses minimum wages applicable to the entire economy and considers all patents. In doing so, this study aims to add a new perspective to the ongoing discussion of wage floors and innovation, hoping to provide some additional understanding in this complex and multi-faceted field.

This thesis is divided into 9 chapters. After the introduction, chapter 2 reviews previous literature on the thesis' research question. The chapter describes both historical, theoretical, and empirical literature on the subject. Chapter 3 describes the institutional background that determines wage floors and innovation, with a particular focus on the United States. Chapter 4 explains the empirical method used for the analysis. The chapter gives a detailed description of the DiD methodology with a two-way fixed effects estimator. I also discuss the identifying assumptions required for the approach, including some specific considerations in my model which are necessary in order to deal with treatment heterogeneity. Chapter 5 describes the data used in the analysis, and chapter 6 shows the main results. Chapter 7 evaluates the statistical validity of my DiD estimates, first by examining pre-trends and, second, by investigating treatment effect heterogeneity. Chapter 8 contains a discussion of the results, particularly as they relate to the theoretical model of the thesis and the external validity of my findings. Finally, in chapter 9 I summarize the thesis, give concluding remarks, and suggest areas of further research.

2 Literature Review

In this chapter, I provide a brief overview of the literature linking minimum wages to innovation. As the literature on this topic is vast, this section is a focused summary of the work that is the most relevant for my empirical analysis. In section 2.1 I present historical proponents of the view that high labor costs stimulate innovation, as well as critics of this view. Section 2.2 concerns Acemoğlu's model of labor scarcity, which is the main theoretical framework for the thesis. Finally, section 2.3 explores empirical evidence.

2.1 Historical Arguments for the View That Labor Scarcity Encourages Innovation

It has long been speculated that there might be a link between the availability of low-wage labor and technological progress. John Hicks raised this possibility in his work *The Theory of Wages* where he argued that "a change in the relative prices of the factors of production is itself a spur to invention, and to invention of a particular kind – directed to economizing the use of a factor which has become relatively expensive" (Hicks, 1932, p. 124).

In an article from 1946 called *Causes of the Superior Efficiency of U.S.A. Industry as Compared with British Industry*, Erwin Rothbarth examines the claim that the superior efficiency of American industry as compared to British industry beginning in the latter part of the 19th century was due to the larger size of the American market. Rothbarth finds this explanation unsatisfactory for a variety of reasons, and instead presents alternative theories for the divide. One theory he briefly touches on, is the theory that abundant land in the early days of the United States' history made it less attractive for laborers to work in manufacturing, forcing the use of capital-intensive production and innovative practices to compensate for the lack of cheap labor. He writes:

> In any country where land is readily available in large quantities, labour is likely to be expensive. For the income of the industrial worker must be sufficiently high to present an attractive alternative to his cultivating the land for his own profit. Thus the high productivity of labour in American industry at the beginning of this century can be explained by the fact that industry had to install labour-saving equipment and to economise in the use of labour until its productivity was sufficiently far higher than it was in agriculture to enable relatively attractive wages to be paid in industry. (Rothbarth, 1946, p. 385)

Habakkuk expanded this idea in his 1962 work *American and British Technology in the Nineteenth Century: The Search for Labor-Saving Inventions* into what is today known as the Habakkuk thesis. The thesis states that labor scarcity (whether caused by land abundance, minimum wages, or other means) can lead to labor-saving technological innovations. In the same way as Rothbarth, Habakkuk justifies his theory with a historical analysis of the development and adoption of labor-saving technologies in the United States and Britain during the 18- and 1900s. Habakkuk argues that the two countries' differing approaches to labor and resource availability shaped their respective technological trajectories. In the United States, land abundance made industrial labor relatively scarce and expensive, which motivated the search for labor-saving inventions. This led to the development of technologies that prioritized efficiency, standardization, and mechanization, which he calls the American system of manufactures (perhaps best exemplified by the Henry Ford assembly line). These innovations allowed the United States to rapidly industrialize and gain a competitive advantage in the global economy. In contrast, Britain had a more abundant and affordable labor supply, which reduced the urgency to adopt labor-saving technologies. As a result, British industry focused on more specialized, craft-based production techniques. This approach allowed for greater product diversity and customization but limited the scale of British industrialization compared to the United States.

More recently, Robert Allen (2009) has argued that high wage costs were fundamental for the development of industrial cotton manufacture and thereby for the Industrial Revolution. According to Allen, high wages in Britain led entrepreneurs and inventors to seek labor-saving technologies and innovations to reduce production costs. This spurred the development of machinery and new production techniques, such as the spinning jenny and the steam engine. In contrast, lower-wage countries had less incentive to invest in labor-saving technologies, which explains why the Industrial Revolution did not occur elsewhere. Adoption of the new technology was much slower in France and India, which Allen suggests was because of the lower wages (and thus the lower incentive to save on labor costs). Related to this, Mark Elvin (1972) claims that a spinning wheel suitable for industrial cotton manufacture was invented in 14th century China but was abandoned because cheap Chinese labor was readily available and made the technology unprofitable.

2.1.1 Criticisms of the View That Labor Scarcity Encourages Innovation

Peter Temin is an economic historian and a prominent critic of Habakkuk. In his 1966 paper *Labor Scarcity and the Problem of American Industrial Efficiency in the 1850's* he labels Habakkuk's sources on American manufactural superiority as anecdotal. Moreover, he claims that the theory assumes that the price of land only affected agriculture, and that land's use in manufacture is left out by the thesis. Of the three traditional factors of production – land, labor, and capital – the Habakkuk thesis only concerns itself with the price of land and labor for agriculture, and labor and capital for manufacture, ignoring a third factor in both productions. According to Temin, if there indeed were differences in the efficiency of the two countries (which he is not willing to concede), then the interest rates of the two countries would be the more important reason for this difference.

For this thesis, the particulars of what factors were the most important in developing the American system of manufactures and how different it was to Britain is less relevant than the broader issue of how labor scarcity relates to innovation. I will therefore concentrate on how the ideas of Habakkuk, Rothbarth, and their likes apply to contemporary issues, not on the accuracy of their analysis of 19th and 20th century America.

While the intuition of the Habakkuk thesis is straightforward, it is not obvious that it is correct when looking at it through the lens of various economic models. In fact, most macroeconomic models reach the opposite conclusion – that high wage costs discourage technological progress.[1] While labor scarcity could theoretically make technological investment relatively more profitable, high wage costs also reduce both the profitability of firms and the availability of workers that may use the new technologies, both of which could discourage technological innovation.

2.2 Acemoğlu's Model of Labor Scarcity

The paper *When Does Labor Scarcity Encourage Innovation?* by Daron Acemoğlu (2010), explores the relationship between labor scarcity and the adoption of new technologies. The central question is under which conditions labor scarcity increases innovation, and when it does not. Though the term "labor scarcity" is used throughout, in a competitive labor market this is equivalent to an exogenous increase in wage costs exemplified by a minimum wage.

Acemoğlu's paper builds on Habakkuk's thesis and extends it through the development of a theoretical model to analyze the conditions under which labor scarcity encourages or discourages innovation. While Habakkuk's work was largely historical and empirical, Acemoğlu uses economic theory to formalize the relationship between labor scarcity and innovation, providing

[1] Ricardo (1951) is an early example of this view.

a more general framework to understand the phenomenon. While this is not the only model of this issue, Acemoğlu's model is comprehensive and meticulous and makes it easy to analyze the impact of a minimum wage increase. Hence, it will be the main theoretical model for this thesis.[2]

In the paper, Acemoğlu introduces the terms "strongly labor-saving" and "strongly labor-complementary" to refer to the different ways in which technologies can interact with labor in the production process. Strongly labor-saving technology refers to technologies that reduce the need for labor in the production process. These technologies are designed to replace human labor with machines or automated processes. When labor is scarce, there's a greater incentive for firms to adopt these labor-saving technologies, as they can help to maintain or increase output levels while reducing the need for labor. Strongly labor-complementary technology, on the other hand, refers to technologies that enhance the marginal productivity of labor. These technologies are designed to work alongside human labor and increase the output that each worker (i.e., its marginal product) can produce. Hence, when these technologies are adopted, the demand for labor increases because each worker is now more productive.

Acemoğlu uses these concepts to analyze how labor scarcity can influence the choice of technologies and the rate of innovation. He suggests that labor scarcity will encourage innovation if technologies are generally labor-saving but will not do so if instead technologies and labor are complementary. The specific direction of technological change therefore depends on what form technologies take, which in turn depends on various factors, including the elasticity of substitution between input factors and the degree to which returns to scale are diminishing.

Acemoğlu also shows that most growth and macroeconomic models either implicitly or explicitly assume that technological advances increase the marginal product of labor, thereby making technology labor-complementary. He further presents plausible environments where he argues that technological change is instead labor-saving.

For simplicity's sake, this thesis will present the version of Acemoğlu's model where there is a single supplier of technology and a static framework, as the more complex versions found in his paper result in similar outcomes as it relates to the areas of interest for this thesis. In this

[2] See James and Skinner (1985) for another model of labor scarcity.

model, technology is represented with a single-dimensional variable, θ. The factors of production are capital K, labor L, and a third factor T which for the sake of this analysis can be land or some other nonlabor factor of production. The model distinguishes between two cases, one in which technology θ "augments" capital, and one where it augments labor (i.e., improves productivity of the factor). I will first tackle capital-augmenting technology before turning my attention towards labor-augmenting technology.

As using a Cobb-Douglas product function always yields labor-complimentary technology, Acemoğlu instead proposes using a constant elasticity of substitution production function[3], where θ represents capital-augmenting technology:

$$G(L,K,T,\theta) = [(1-\eta)(\theta K)^{\frac{\sigma-1}{\sigma}} + \eta L^{\frac{\sigma-1}{\sigma}}]^{\frac{\gamma\sigma}{\sigma-1}} \times T^{1-\gamma},$$

where $\eta \in (0,1)$ and $\gamma \in (0,1)$ for the function G. σ shows the elasticity of substitution between labor and capital, η is the labor and capital shares, and γ captures the returns to scale so that when $\gamma < 1$, there are decreasing returns to labor and capital holding land constant.

It follows from G that, for capital-augmenting technology to be strongly labor-saving, it is necessary for $\gamma < 1$ and for $\sigma > 1$ to such a degree that

$$1-\gamma > \frac{1}{\sigma}$$

is satisfied.[4] The main takeaway from this result is that capital-augmenting technology requires diminishing returns to scale and sufficiently large elasticities of substitution between capital and labor in order to be labor-saving. According to the model, in situations where it is easy to exchange labor for capital, and where scaling up production gives decreasing returns, increased labor scarcity will make capital-augmenting technology more attractive in order to save on labor costs.

[3] An alternative to the constant elasticity of substitution production function is a task-based production function in which technological changes are not factor-augmenting, but where instead "tasks" are either automated or still produced by labor. See Acemoğlu and Autor (2011).
[4] See Acemoğlu (2010) for more detailed calculations.

Let us now turn to technology which makes labor more productive. The G function is now expressed like so:

$$G(L,K,T,\theta) = [(1-\eta)K^{\frac{\sigma-1}{\sigma}} + \eta(\theta L)^{\frac{\sigma-1}{\sigma}}]^{\frac{\gamma\sigma}{\sigma-1}} \times T^{1-\gamma}.$$

We can define labor's share of return to capital as

$$S_L \equiv \frac{wL}{rK} = \frac{\eta(\theta L)^{\frac{\sigma-1}{\sigma}}}{(1-\eta)K^{\frac{\sigma-1}{\sigma}}} > 0,$$

where w is the wage paid to workers and r is the interest paid to owners of capital (the marginal product of the factors). For technology to be labor saving requires

$$S_L < \frac{1-\sigma}{\sigma\gamma}$$

when technology is labor-augmenting. As with capital-augmenting technology, strong diminishing returns to scale makes it more likely that technology will be labor-saving, however it is now not a *requirement* that the returns to scale are diminishing, as long as $\gamma < \frac{1-\sigma}{\sigma S_L}$. What *is* necessary however, is that $\sigma < 1$, which is the opposite of the requirement when technology augments capital. For technology to replace labor when it makes capital more productive, it must be easy to switch from labor to capital, but when technology makes labor more productive it must instead be hard to switch away from capital for less labor to be employed by the new technology.

Looking at the extremes can help to illustrate under what circumstances exogenous wage increases lead to more labor-saving technology. If labor becomes more expensive, technology improves the returns from capital, and it is trivial to switch between capital and labor, then firms will become more capital-intensive and invest more in capital-augmenting technology. If instead technology improves labor's productivity, but it is impossible to switch to a different capital-labor ratio, then increased labor costs, i.e., less laborers employed, will necessarily bring with it a proportional reduction in the capital; employers cannot shift away from the labor costs by employing more capital. Originally, when the workers earned the competitive wage, adopting the labor-augmenting technology would make the workers more productive but also more expensive as the wage in a competitive market is the worker's marginal productivity. However, if wages of the worker have increased exogenously – without an accompanying productivity increase – then technology can boost their productivity without increasing wage costs (more than wages have already increased). In both examples strong diminishing returns to scale strengthens the case for shifting away from workers and towards technological investments because, when the first few units of a factor are significantly more productive than the later ones, the workers on the margin generate modest revenue anyways, and a productivity boost by instead investing in technology becomes relatively more attractive.

There are three cases where we would not expect increased labor costs to lead to more innovation (given constant or weakly diminishing returns to scale). The first case is if technology improves the returns from capital, but it is hard to switch between the factors of production. In this case, a reduction in labor would pull capital down with it, and increased productivity from capital would not be any more attractive than before. The second case is if technology improves labor but it is also easy to switch towards more capital use. Employers can in that case avoid all the losses from increased labor costs and will not have to invest in new labor augmenting technology. The third case is if the wage is not competitive. In this case, increased wage costs will not result in a reduction of any inputs.

In summary, the effect of an exogenous increase in labor costs on innovation depends on the particulars of the situation. Technology is always more labor-saving the more the returns to scale are diminishing. Additionally, if technology mostly improves the productivity of *capital*, then the elasticity of substitution between the factors must be large. However, if technology mostly improves the productivity of *workers*, then the elasticity must be low for technology to be labor-saving.

2.3 Empirical Literature

Recent studies suggest that raising the minimum wage could result in an increase in technological innovations. Next, I will describe some of the studies showing this effect.

The most widely studied aspect of the minimum wage is its effect on employment. One such study, which has relevance for this thesis, is the paper *People Versus Machines: The Impact of*

Minimum Wages on Automatable Jobs by Grace Lordan and David Neumark (2018). The paper looks at minimum wages in the United States in the years 1980 to 2015. The findings indicate that, following a minimum wage increase, workers who have easily automatable jobs (i.e., jobs where it is relatively easy for the employer to substitute labor for machines) experience an increased risk of becoming unemployed. Whether the shift is caused by adoption or by development of new technologies is outside the scope of the study, but it does establish a link between minimum wages and new technologies.

Dechezleprêtre et al. (2019) explores the impact of a rise in low-skill labor costs on automation innovation by examining how automation-related keywords in patent texts relate to spontaneous changes in wage levels. The reliability of this approach is validated by its correlation with a decline in routine tasks. The study takes advantage of differences in companies' exposure to global markets to isolate exogenous wage changes in 41 countries in the period 1997-2011. The authors observe that increases in low-skill wages, prompted by international market forces, stimulate more automation innovation. They estimate the elasticity to be between 2 and 4. Conversely, they report a reverse effect for high-skill wages.

The paper also investigates the effects of Germany's Hartz reform, which lowered labor costs by enhancing labor market flexibility and reducing social contributions for workers. The study shows that foreign firms that were highly exposed to Germany decreased their automation innovation relative to other foreign companies post-reform, aligning with the study's other findings.

The research of Dechezleprêtre et al. is primarily concentrated on the manufacturing industry, contrasting with my thesis which is more relevant to sectors where the minimum wage has a significant impact (primarily the service sector). In the manufacturing industry, the minimum wage is a poor estimate of an exogenous wage increase as manufacturing wages are generally well above the minimum wage (Statista Research Department, 2023).

Like my thesis, Amrita Nain and Yan Wang (2021) exploit changes in the minimum wage in the United States to study how higher costs for unskilled labor stimulate greater labor-saving innovation. Similarly to Dechezleprêtre et al., Nain and Wang uses patent applications with keywords relating to automation as their dependent variable, as well as citations of these patents. The paper uses an event study approach to compare automation patents in a two-year period before and after the introduction of a higher minimum wage in the years 1987-2017. Their analysis finds a correlation between increases in the minimum wage and greater numbers of automation patent applications, along with more citations of such patents. The impacts are more pronounced in states where the wage increases are more binding.

Some studies examine specific products or industries to show that an increase in labor costs lead to more labor-saving technology. Hannan and McDowell (1984) show that labor costs played a role in how quick banks were to adopt ATMs (automatic teller machines). Acemoğlu and Finkelstein (2008) describe how increased wages for health care workers, caused by the Medicare Prospective Payment System reform, resulted in a shift towards a more capital-intensive operation and moreover to the adoption of a range of new medical technologies. Manuelli and Seshardi (2014) argue that tractors were only utilized in the US agricultural sector after wages increased to such an extent that tractors became more profitable than labor-intensive horses. Clemens et al. (2018) show that employers responded by using more technological production techniques instead of hiring domestic workers after the exclusion of half a million Mexican farmworkers from the United States in the 1960s.

3 Institutional Background: Wage Floors and Innovation in the United States

3.1 Minimum Wages and Labor Unions

A wage floor is an economic term referring to the lowest allowable wage that an employer can pay its employees. When this floor lies above the equilibrium wage that would be given in an entirely free market, employers must make a decision on whether or not it is worthwhile to keep employing the same number of workers at the higher wage. Wage floors are often a subject of debate among economists and policymakers. Advocates argue that they help reduce poverty and income inequality by ensuring that workers earn a basic living wage. On the other hand, opponents contend that imposing wage floors can have negative consequences, such as increased unemployment and higher prices for goods and services, as businesses may pass on the higher labor costs to consumers. Empirical work on these questions have given varying results, and there is no consensus in the field on the general effects of a wage floor (Neumark, 2017).

Generally, there are two main ways of establishing a wage floor: with a minimum wage or through collective bargaining.[5] A minimum wage is a legally mandated wage level set by the government, below which employers cannot pay their employees. The main objective of minimum wage laws is to ensure that workers receive a basic level of income. The government sets the wage floor, and it is uniformly applied to all eligible workers regardless of the industry or occupation, serving as a baseline for worker compensation. In most countries, the minimum wage is updated periodically to account for inflation and changes in living costs, although this is not the case in the United States. A wage floor can also be set by labor unions through collective bargaining. Labor unions are organized groups of workers that bargain with employers for better wages, working conditions, and other benefits. Unions provide a wage floor by negotiating contracts (also known as collective bargaining agreements) with employers on behalf of their members. These contracts specify the wage floors and other benefits that union members are entitled to receive.

While both legally mandated minimum wages and labor unions aim to provide wage floors for workers, they differ in several important ways. First, minimum wage laws apply to all eligible workers within a given jurisdiction, while union-negotiated wage floors are usually applicable only to members of the respective labor union.[6] Second, government-set minimum wage laws are generally uniform across industries and occupations, whereas union-negotiated wage floors can be tailored to suit the specific needs and conditions of a particular industry or occupation. Third, labor unions negotiate for additional benefits, such as health insurance, retirement plans, and paid time off, while minimum wage laws only address the base wage rate. Finally, labor unions provide workers with collective bargaining power when negotiating with employers, which can lead to higher wages and better working conditions than might be achieved individually. In contrast, minimum wage laws are determined by the government and do not involve direct negotiation between workers and employers.

In the United States, both labor unions and minimum wage laws contribute to establishing wage floors for workers. However, compared to many other developed countries, minimum wages

---

[5] Although some countries also utilize wage councils, worker cooperatives and sectoral bargaining to reach the same goals as the above mentioned, this is virtually unused in the United States and will not be discussed in this thesis. See Slinn (2023) for a discussion on these approaches.

[6] However, mandatory extensions are sometimes used outside of the United States to make the union wage the effective minimum wage of the sector (Arpaia et al., 2017).

play a proportionately larger role than labor unions. In recent decades, union membership in the United States has declined, reducing the overall impact of unions on wages. As of 2022, the Bureau of Labor Statistics reported that only 10.1% of US workers were union members. Additionally, union representation is more concentrated in specific industries, such as public sector jobs, construction, transportation, and education, while it is lower in some private sectors, such as retail and service industries (Bureau of Labor Statistics, 2023b). This disparity means that the impact of labor unions on wages is not present in large sections of the US economy, and that minimum wage laws have a broader impact because they cover all eligible workers, irrespective of union membership or industry.

3.1.1 Federal and State Minimum Wages

The concept of a federal minimum wage was first introduced in the United States with the Fair Labor Standards Act of 1938. The federal minimum wage, as of 2023, is at 7.25 USD per hour and was last raised in 2009 (Fair Labor Standards Act, § 206). In addition to the federal government, individual states, and even some cities, have the authority to set their own minimum wage rates, which can be higher than the federal rate. Employers are required to pay their employees the highest of the applicable federal, state, or local minimum wage.

Minimum wage laws apply to most workers in the United States, with some exceptions. For example, tipped employees, like restaurant servers, can be paid a lower minimum wage ($2.13 per hour at the federal level), provided that their tips bring their total earnings up to at least the applicable minimum wage. There are also exemptions for certain types of workers, such as seasonal employees, farmworkers, and full-time students working in certain sectors.

One important feature of American minimum wages, whether they be federal or local, is that they are not automatically adjusted for inflation. Instead, the minimum wage is only sporadically updated whenever the state or the federal government passes new laws requiring a higher minimum wage. This means that, in periods where the minimum wage is not adjusted, a general increase in prices reduces the real value of the minimum wage (in terms of what it can buy) at the same yearly rate as the inflation rate. The United States is distinct in this regard, as, in most other countries, minimum wages are either automatically adjusted or periodically reviewed (Desilver, 2021).

In my empirical analysis, I will exploit two sources of variation in minimum wages across states in order to provide causal evidence on the effect of minimum wages on innovation: The first is changes in the state minimum wages; the second stems from the fact that US minimum wages are not inflation adjusted. The lack of adjustment for inflation, combined with the fact that the different states to a large extent set their own minimum wages, results in good prerequisites for comparisons between the states of the effects of having different levels of minimum wages. Because inflation gradually erodes the purchasing power of the minimum wage until it promptly rises again when a new law is passed, minimum wage raises in the United States become "events" which can be studied. This would not be possible if instead the minimum wage increased slowly and smoothly. Importantly, not all states raise their wages at the same time which enables us to use states as controls for each other. I will go into more detail about my empirical method in chapter 4.1.

3.2 Innovation

Innovation plays a vital role in driving economic growth, fostering job creation, and improving the overall standard of living. Innovation will in this thesis refer to the implementation and diffusion of new or improved products, services, processes, or business models that create value and enhance productivity or efficiency. Innovation can stem from advancements in technology, research, knowledge, or the creative application of existing resources and ideas.

Innovation can be categorized in many different ways, but the main focus will here be on product innovation, process innovation, and technological innovation. Product innovation refers to the introduction of new or significantly improved goods or services that meet consumer needs, offer better quality or performance, or create new market opportunities. Process innovation is the development of new or more efficient methods of production, distribution, or service delivery that reduce costs, increase productivity, or improve quality. Lastly, technological innovation means the application of new or improved technologies or scientific discoveries that lead to the creation of novel products, services, or processes.

In the United States, the process for patenting an innovation is overseen by the United States Patent and Trademark Office. A patent gives an inventor the exclusive rights to make, use, sell, and distribute their invention for a limited period of time – in the case of the United States: for 20 years. A patent application typically includes a written document outlining the invention, a set of claims defining the invention, an oath or declaration, and diagrams or drawings, if necessary. A patent application can be provisional at first, which provides the means to establish an early effective filing date in a later filed non-provisional patent application and allows the term "Patent Pending" to be applied before a patent examiner has granted the patent. The examiner will determine whether the invention is novel, non-obvious, and useful, in which case the patent will be granted. Maintenance fees must be paid at 3.5, 7.5, and 11.5 years after the patent is granted to keep it in force (United States Patent and Trademark Office, 2023).

Not all innovations are patentable. Generally, most forms of product innovations and technological innovations can be patented as long as they fulfil the requirements of being novel, non-obvious, and useful. However, the patentability of process innovation is more uncertain. If the process innovation is a purely mental process or methods of doing business, it could fall under "abstract ideas", which cannot be patented (United States Patent and Trademark Office, 2023). As such, using patents as an indicator of innovation could mean that certain process innovations are not captured. Additionally, while both the development and the adoption of new technologies are generally categorized as innovation, patents only encapsulate the effect of the former.

On a general basis, the United States is a highly innovative country, overtaking Sweden to reach the 2nd place in the Global Innovation Index report of 2022, surpassed only by Switzerland (World Intellectual Property Organization, 2022). The report credits this achievement to the country's robust ecosystem of universities, research institutions, and tech companies, which play a vital role in fostering innovation. America's strong intellectual property rights, an economic climate conducive to entrepreneurship, and substantial investment in research and development all contribute to its high ranking.

Despite the general high degree of innovation, in some areas the United States lag behind other developed economies. Mika Pajarinen, Petri Rouvinen, and Anders Ekeland (2015) show that the United States have more workers whose jobs are at risk of being automated than Norway and Finland do. Erling Barth, Karl O. Moene, and Fredrik Willumsen (2014) indicate that this could be because private investments have made low-income labor more productive in Scandinavia than in the United States. They also show that all the Scandinavian countries have historically had higher growth rates than the United States. Fölster (2018) highlights that many low-skilled jobs that are still done by humans in the US have already been automated in Norway, but that new jobs were also created at a similar rate to the old jobs that were automated. The empirical analysis of this thesis can indirectly shed some light on whether the higher wages of low-skilled workers in Scandinavian countries, owing to the strong presence of unions, could be the cause of this difference.

4 Empirical Analysis

4.1 Identification Strategy

In a random experiment, isolating the causal relationship is done by randomly assigning units into either a treatment group or a control group, ensuring the two groups to be virtually identical except for the treatment. However, when studying most economic questions, this is not viable. Instead, I use a DiD estimation technique to isolate a causal relationship between a wage floor and innovation. DiD estimation is a statistical technique which can be used to evaluate the effect of an intervention (like the raising of a minimum wage) on an outcome variable by comparing the differences in outcomes for a treated group and a control group before and after the policy introduction. However, as it is impossible to randomly assign states into control and treatment groups, DiD instead uses quasi-natural experiments – where a treatment naturally occurs in one group, but not in another similar one. Still, if the two groups differ in other aspects than the treatment, this is something which must be corrected for. A DiD approach therefore takes into account that the groups could have already been dissimilar before the treatment occurred and assumes that the difference between them would have remained the same in the absence of the treatment, i.e., the trends of their dependent variable would have been parallel.

Normally, an omitted variable might cause the regression to be biased if the groups self-select whether to receive treatment. This is the case if the group that chose to introduce the treatment possesses characteristics which are different to the group that did not, and those characteristics are correlated with the dependent variable. The minimum wage level set by different states is clearly not a random choice, it is a policy decision which could potentially be correlated with various economic outcomes, including the level of innovation in the different states. My empirical strategy therefore needs to consider the self-selection bias present in the treatment group, i.e., the states that raise the minimum wage.

It should be noted that innovation is typically not a primary concern for policymakers when deciding on the minimum wage level. Other effects, like the effect on the unemployment rate, have been more thoroughly studied, and so we would expect there to be a higher level of self-selection based on these (better known) outcomes than on the degree of innovation. Also, to the extent that the different minimum wage level is caused by a difference in timing, self-selection bias is less of a concern. Because inflation gradually lowers the real value of the minimum wage, the difference between when the states raise their minimum wage could mostly be a timing difference in how long it has been since last time they raised it. If the distinction between wages primarily is a timing difference of when to raise it, rather than a difference in policy preference between the states, then one would not expect there to be systematic differences between the treatment and the control group. However, to the extent that the opposite is true, self-selection will still be a problem.

DiD deals with this problem by comparing the change in outcomes over time between the treatment and control groups to control for time-invariant unobserved characteristics. The sample might still suffer from self-selection bias, but this is less of a problem than it would otherwise be when using DiD. Because we are looking at trends, the assumption that the two states are identical but for the independent variable, is not required. One only requires their relative trends to be similar, which is a weaker assumption, and one which I will discuss further in chapter 4.2.

4.1.1 The Two-Way Fixed Effects Estimator

The particular form of DiD regression used in this paper is a two-way fixed effects (TWFE) estimator. This is a form of DiD regression which uses panel data (data that consists of multiple observations over time for a set of units – in this case US states) to control for both time-invariant unobserved characteristics of the cross-sectional units (e.g., unit-specific effects) and time-specific factors that affect all units in the panel (e.g., time-specific effects). Doing so, we can allow for self-selection, so long as the trends of the groups are the same.

The main purpose of a TWFE regression is to isolate the relationship between an independent variable and a dependent variable while accounting for unobservable factors that are either time fixed or unit fixed, and that may be correlated with both the dependent and independent variables and their trends. This helps to mitigate potential omitted variable bias and tease out the causal effect of the independent variable on the dependent variable.

The TWFE functional form can be written as:

$$Y_{i(t+2)} = \alpha_i + \alpha_t + \beta_1 D_{it} + \beta_2 X'_{it} + \epsilon_{it}$$

The outcome $Y_{i(t+2)}$ is dependent on the state fixed effects α_i, the time fixed effects α_t, the treatment effect β_1, the treatment level (dosage) D_{it}, covariates X_{it}, the vector of coefficients for each covariate β_2, and the error term ϵ_{it}. On average the error term should be zero, and β_1 should capture the average treatment effect (ATT), which is our main parameter of interest. For my analysis, the outcome variable $Y_{i(t+2)}$ is patents granted in each state, and the treatment D_{it} is the real state minimum wage. Because patents take time to develop, there is a likely a delay before the effect of a higher minimum wage manifests itself in patent registrations. The outcome variable is therefore shifted two years back, so that the treatment effect is measured on patents granted two years after the wage raise.[7] The subscript $(t + 2)$ is used to reflect this.

Various control variables in X_{it} are included in the regression in order to adjust for any differences in the trends between the groups caused by these. Included controls are: population, how urban the state is, how educated the state is, state GDP, how big the service, manufacturing, and agricultural sector of the state economy are, how prevalent low-income households are in the state, how the state leans politically, and what political party rules the state.

The mentioned controls are included to account for the possibility that they are confounders, i.e., variables that cause a spurious association between the treatment and the outcome variable. I will discuss briefly how each control variable could be confounding. Keep in mind that control variables only control for shocks in confounding factors; time-invariant state characteristics, and state-invariant general trends are captured, respectively by the state and time fixed effects.

Population: A population increase might be correlated with minimum wage increases due to the fact that the politics of larger states differentiate themselves from the political decisions of smaller states (Abbott & Levine, 1991). Population can cause more patents to be developed simply by there being more potential inventors, as well as the potential for there to be scalar effects of a larger population on innovation.

[7] Two years was chosen as the appropriate length of time following evidence from de Rassenfosse (2013)

Urbanization: How urban a state is has the potential to affect patent development if a disproportionate amount of patents are developed in cities, which there is some evidence for in the United States, though not in most other countries (Fritsch & Wyrwich, 2021). The urbanization rate can be linked to the minimum wage because of differing wage patterns in cities and in the countryside (Chin, 1998).

Education: As educated individuals tend to have a higher wage, the pressure to raise the minimum wage could be lower from states where there are more individuals with a higher education. Education can also cause more patents to be developed if educated individuals innovate more.

Gross domestic product: A wealthier state might invest more in innovation than a poor state. How wealthy a state is could also be related to whether there is a high or a low minimum wage in the state.

Industry size: The minimum wage is not equally applicable to all sectors of the economy. Leisure and hospitality is by far the industry with the most minimum wage workers (Statista Research Department, 2023). To account for this, the size of the service sector is included in the analysis. The agricultural sector is often exempt from minimum wage laws, and as such it is also included as a control variable. In addition, the decline in manufacturing jobs have been an important factor shaping various outcomes in many states, and it is therefore also controlled for.

Low income: The prevalence of low-income individuals in a state is likely to correlate with the amount of people that are dependent on the minimum wage in the state. The political preferences of these citizens when it comes to the minimum wage could help shape minimum wage legislation. It could also affect patent production if poorer persons are less likely to register patents.

Political preferences: As the minimum wage rate is a political choice, the political leanings of the populace will undoubtably play into its level. If there is also a difference between Democrat and Republican voters in terms of likelihood to register patents, then this can be a factor to take into consideration as well.

Political control: Similarly to the political leanings of the population, what political party controls the state will likely correlate with the level of the minimum wage. Politicians can also shape the economic environment to be more or less conducive to innovations.

Some of these variables are expressed in logarithmic form to reflect that we are not interested in the absolute numbers of the change, but in the ratio of change (this is reflected by their linear histograms, which are not normally distributed). In particular, the population, GDP, and patent variables are log transformed.

4.2 Identifying Assumptions

For a TWFE DiD regression to accurately measure the causal effect of the treatment variable on the dependent variable, certain underlying assumptions must be fulfilled. Next, I will describe these assumptions one by one and discuss whether it is reasonable to assume that they hold for my analysis.

Positivity

This positivity assumption is the assumption that units receive different levels of the treatment variable. If this assumption does not hold and all units have the same level of treatment, then we have no basis for comparison between the units. In our case, the different states set their minimum wage level differently, and so we are able to compare outcomes. One potential problem is that no states in the sample have a minimum wage at zero as no state can go below the federal minimum wage. To get around this, the treatment variable will be continuous in the model specification, allowing for differentiation between the states.

Parallel trends

One very important prerequisite for DiD is parallel trends, which means that in the absence of the treatment, the average outcomes for the treatment and control groups would have followed the same trend over time. This assumption is key to establish the counterfactual – what would have happened to the treatment group if they had not received the treatment. I will dedicate substantial space in my thesis to investigate how this assumption holds up; see section 4.2.2 and 7.1 for more details.

No time-varying confounders

Related to parallel trends, DiD also assumes that there is no other event or shock that coincides with the treatment and differentially affects the treatment and control groups. Although DiD automatically controls for time-invariant factors, there could still be time-variant shocks that affected one group (treatment or control), but not the other. This difference would then be incorrectly attributed to the treatment effect and bias the results. To account for this, I will be controlling for the time-varying confounders described above.

No reverse causality

Using DiD requires that the dependent variable is exogenous such that it does not causally affect the independent variables (treatment and confounders). If the opposite is the case, we have reverse causality and biased results. Because in my case there is a delayed reaction in the dependent variable (patents), we know for sure that this is not an issue in our specification. It is impossible for the amount of patents granted to causally affect what happened two years prior with the independent variables. For future outcomes on the other hand, it is likely that patents have an effect (particularly on confounders such as GDP), but patents are not able to affect outcomes in the past. The dependent variable is therefore probably not strictly exogenous to all independent variables, but it is sequentially exogenous, which is enough for the assumption to hold.

Stable unit treatment value

The stable unit treatment value assumption (SUTVA for short) is the assumption that the treatment is consistent between all units that receive it, and that units are unaffected by other units receiving treatment. The latter part of this assumption could potentially be a problem in my study. It could be that as innovative business practices arise because of the minimum wage in one state, it spreads to a neighboring state which did not raise its minimum wage. In addition, people can move across state borders during the observation period, so the composition of the treatment and control groups is not necessarily stable (although citizens must follow the laws of their new state of residence). Because of this weakness in the analysis, the effect could potentially be larger than the main results show.

Besides the standard assumptions required for a TWFE DiD regression, there are some additional considerations required when the treatment variable is continuous and have a staggered adoption. I will discuss these restrictions in the next section.

4.2.1 Staggered Treatment Timing and Continuous Treatment

Unlike the "standard" DiD approach, the changes in treatment level do not happen simultaneously in all units in my sample. Because of this, I must make an additional assumption when using the TWFE estimator, which is that the treatment effect is homogenous (Callaway & Sant'Anna, 2021). What this means is that the effectiveness of the treatment variable on the dependent variable must not diverge across units or across time. The strict interpretation of this assumption – that there is no heterogeneity in how minimum wages affect innovation in different states and over time – is a rather strong one and is not likely to be wholly accurate. To account for this, and to identify the extent of the problem, I will explore alternative estimators to the TWFE estimator that do not require homogenous treatment effects in chapter 7.2.

Another distinction between this thesis' specification and the textbook DiD, is that my regression model uses a continuous treatment, not a binary one. This has consequences for the parallel trends assumption, which I will examine in detail in the next subsection (chapter 4.2.2). The reason why I nevertheless keep the treatment continuous for the main analysis instead of discretizing it into a dummy variable, is multifold. One reason is that having a continuous treatment allows one to run a regression even when no units remain untreated throughout the period (which is the case for US states' minimum wage level) by comparing different treatment levels to each other. If instead the minimum wage level is to be discretized, the classification would have to be a dummy variable where 0 meant "at the federal level" and 1 meant "somewhere above the federal level". Doing so would make it unclear what a treatment level of 1 means, as the states in this category can vary considerably in their actual level of minimum wage. In addition, being "untreated" would have a wavering meaning, as the real federal minimum wage changes from year to year. Finally, on a more fundamental level, there is no getting around assuming some basic facts about how a local effect generalizes to other cases, no matter what econometrical methods one uses. Essentially, all research designs are various ways of getting good estimates of the local effect one is investigating, but if the area of interest is not just in the specific example but in the general effect of the treatment, one must necessarily assume that other manifestations of the phenomenon share some characteristics with the observed events. Not classifying the minimum wage as a continuous treatment, when it in fact *is* continuous, would in my opinion just avoid an honest discussion on the assumptions required for the results to generalize.

With this specification, including staggered adoption and a continuous treatment, the TWFE estimator compares differences in patent production before and after changes in states' minimum wage levels with the other states' trends as the counterfactual to find out how various sized jumps in the minimum wage affect the development of patents.

4.2.2 The "Strong" Parallel Trends Assumption

The central underpinning of any DiD analysis is the parallel trends assumption. In the standard DiD case with a binary treatment, the parallel trends assumption is that the control group's trend works as a counterfactual to the trend that the treatment group would have experienced had they not received the treatment. There might be some self-selection into the two groups, but this is largely unproblematic if the trends of their dependent variable follow each other (something which one can investigate with an event study or a parallel trends plot). However, in their paper *Difference-in-Differences with a Continuous Treatment*, Callaway et al. shows that when the treatment is continuous, the regular parallel trends assumption is not enough. The paper argues that one needs a stronger version of the assumption to avoid a biased estimator when using a continuous treatment. The rest of this chapter draws heavily from the work of Callaway et al. (2021).

Unlike with a binary treatment, we need to distinguish between the effect of an increase in treatment and the effect of receiving treatment at all when the treatment is non-binary. We therefore need to define a couple of additional parameters of interest when using a continuous treatment. The average treatment effect on the treated is defined thusly:

$$ATT(a|b) = \mathbb{E}[Y(a) - Y(0)|D = b).$$

The average treatment effect $ATT(a|b)$ is the effect that receiving "dose" (treatment level) a would have on those that actually received dose b. This is not a relationship that is directly observed in the data, only ATTs of the form $ATT(d|d)$ are, meaning the treatment effect of receiving dose d on those who actually receive that dose. When using a binary treatment, all observations are of the form $ATT(d|d)$, because the only dose of interest is the one that everyone in the treatment group receives. However, with a continuous treatment, we observe a range of different dosage levels, and so we must use observations of different dosage levels than the ones units actually experienced as counterfactuals for what they would have experienced had they experienced the same dose. This brings with it a different assumption than in the textbook DiD case. Using the effect on one set of units as counterfactuals for another set of units seems similar to what is already done in all DiD regressions, and indeed it is. This is essentially the parallel trends assumption, but Callaway et al. shows that there is one crucial difference: If there is self-selection into the different dosage levels based on different effectiveness, this will result in bias, and this bias cannot be detected in pre-trend parallel trends graphs or event studies.

Figure 1: Possible ATT-Curves

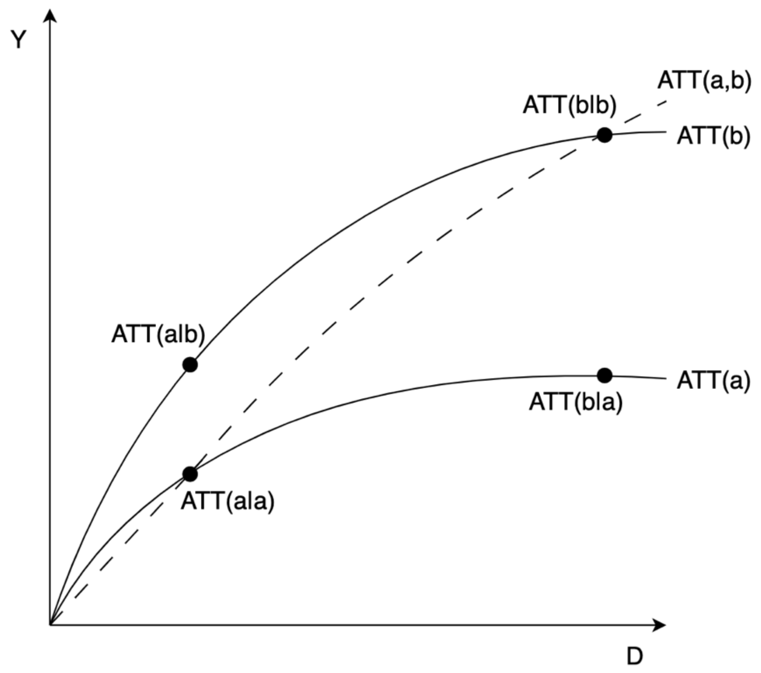

Figure 1 shows the effect that different magnitudes, or doses, of a treatment d has on an outcome Y for two groups – one receiving dose a and one receiving dose b – where $a < b$. We only observe two values: the average treatment effect of receiving a dose of size a on the treated group that received dose a, denoted by $ATT(a|a)$, and the ATT of receiving dose b on the group receiving dose b ($ATT(b|b)$). We can see that the group receiving dose b is both receiving a larger dose of the treatment and experience a larger effect Y than the group receiving dose a is. However, when trying to use the two observed points to guess what the effects on the two groups' Y would be had they experienced different doses than they do, we run into problems. One possibility is that treatment affects them both the same way, in which case if the group receiving b where to receive the other group's dose they would be at point $ATT(a|a)$, just like the group receiving a. This would mean that both groups' ATT-curve is equal to the dotted line in the graph. Another possibility however, is that the two groups have different ATT-curves, and that if the group receiving b were to receive the lower dose of a, they would have an $ATT = ATT(a|b)$.[8]

This is when one would, in a binary DiD, refer to pre-trend graphs and tests to show that the groups are good counterfactuals to each other, but while one must still show that a continuous DiD satisfies the normal parallel trend assumption, this cannot show that the groups have the same ATT-curve, as they will both look the same in the absence of treatment. If one is to use a continuous DiD regression, one has to show evidence (besides pre-trend graphs) suggesting that Callaway et al.'s "strong" parallel trends are satisfied, that is: that there is no systematic connection between how effective the treatment is on the groups and what dose they receive of it.

The question then becomes whether it is a reasonable assumption that the relationship between states' minimum wages and patent production have strong parallel trends. For this to be the case, there must be no state level difference between the effect that minimum wage has on patent production corresponding to the different levels of minimum wage. Note that strong parallel trends do not require that there is *no* link between patent production and minimum wages, just that there is no systematic difference between the *responsiveness* on the number of patents based on the level of minimum wage in the different states.

I will posit that it is not unreasonable to assume strong parallel trends in the case explored by this thesis. While the minimum wage set by each state is obviously not random, there is not much evidence that it has got anything to do with how patent production in the states is affected by wage costs. The different states probably do not choose their dosage level of minimum wage based on expected gains in patents; on the contrary, patents are likely not a concern for the legislature when setting the minimum wage.

[8] To illustrate the problem, consider a fictive study on the effectiveness of pain killers like aspirin. It is reasonable to assume that, to compensate, persons who have a weaker reaction to aspirin might take more pills than people who experience a strong effect of it. One would therefore not expect the size of the dose and the ATT-curves of the groups to be unrelated. A researcher observing individuals choosing different doses of aspirin could not then assume that the effect experienced by high-dose individuals is an adequate counterfactual for the effect that low-dose individuals would have experienced with the same dose.

Moreover, another feature of my dataset is that all units have a non-zero treatment level during all periods. While one can normally not observe strong parallel trends in pre-treatment graphs, the "pre-treatment" that I will check for is actually not the interval before any states are treated but is the interval where they all share the same treatment level. However, this does not mean that the treatment level stays constant during the "pre-treatment", it fluctuates, just at the same rate for all states. If the dose-response relationship between wages and patents was significantly different from state to state, then we would expect their trends to differ at high- and low-points of the real minimum wage's value. However, as we shall see, this is not the case (see figure 2, and the following discussion in section 7.1).

5 Data

This chapter contains descriptions of the data that I collected for the analysis and its sources. All data is in panel form and range from years 1983 to 2015 (the dataset is balanced). This range was primarily chosen because of data availability.

5.1 Treatment and Outcome Variables

This thesis utilizes changes in the minimum wage as the treatment variable for the DiD regression. I compiled the data on the yearly nominal rate of state minimum wages from the US Department of Labor's website (United States Department of Labor, 2023). Some states in the original dataset either lack a minimum wage or have a minimum wage below the federal minimum wage in certain years, however, in these cases the actual minimum wage was the minimum wage set by the federal government. In these instances, I have replaced the state minimum wage by the federal minimum wage in my dataset to show the effective minimum wage. Not all states include agricultural workers in their minimum wage laws. For an extensive list on which states do, see National Agricultural Law Center (2022). In Minnesota, Montana, Ohio, and Oklahoma, small businesses are exempt from the higher state wage. Wage rates are for January 1 of each year.

To get the inflation adjusted minimum wage rate, all years are adjusted by the average value of the Consumer Price Index (Bureau of Labor Statistics, 2023a) in the year. This gives the real effective minimum wage rate in 2020 dollars.

Table 1 shows the variation of when the states first raise their minimum wage above the federal level and enter the treatment group.[9]

Table 1: Year of Treatment Entry

State	Treatment		State	Treatment		State	Treatment
Alabama	Never		Louisiana	Never		Ohio	2007
Alaska	Always		Maine	1988		Oklahoma	Never
Arizona	2007		Maryland	2007		Oregon	1991
Arkansas	2007		Massachusetts	1988		Pennsylvania	2007
California	1991		Michigan	2007		Rhode Island	1988
Colorado	2007		Minnesota	1988		South Carolina	Never
Connecticut	Always		Mississippi	Never		South Dakota	2015
Delaware	1996		Missouri	2007		Tennessee	Never
Florida	2006		Montana	2014		Texas	Never
Georgia	Never		Nebraska	2015		Utah	Never
Hawaii	1988		Nevada	2007		Vermont	1988
Idaho	Never		New Hampshire	1988		Virginia	Never
Illinois	2004		New Jersey	1994		Washington	1991
Indiana	Never		New Mexico	2008		West Virginia	2007
Iowa	1991		New York	2005		Wisconsin	2006
Kansas	Never		North Carolina	2007		Wyoming	Never
Kentucky	Never		North Dakota	Never			

The dependent variable, or the outcome variable, is the number of patents granted in each state by year of grant. This data comes from the US Patent and Trademark Office's Technology Assessment and Forecast database (United States Patent and Trademark Office, 2015). The origin of the patent is determined by the inventor's state of residence, not by where the patent was filed. This is done to avoid strategic choice based on geographical institutional differences in the registration of patents affecting the outcome variable.

[9] By "treatment group" I here mean the states receiving large doses of the treatment relative to the "control group", but technically all states are treated to some degree.

5.2 Control Variables

I collected information on several control variables from different sources. As explained above, these control variables are included in the analysis to account for confounding trends not captured by the fixed effects. These are listed below with a short description.

Population

How the population of each state changes over time is included as a control variable. The source of the data is the US Census Bureau, which estimate the population of all states on July 1st each year (US Census Bureau, 2023a). The population includes everyone who usually reside in the state, not the state of legal residence which is determined differently by each state. Because of this, it does not count US Armed Forces deployed abroad and US citizens living outside of the United States.

Urbanization rate

The US Census Bureau also records the urban percentage of each state, meaning the percentage of the state living in densely populated areas with specific population thresholds. However, this is not recorded annually by the bureau, but every tenth year in the Decennial Census. For the missing years, I have estimated the values by linearly interpolating between the years where there is available data.

Educational attainment

To get a measure of how the education level has evolved in different states over time, I use the percent of the state populations (25 years or older) with a college degree or higher. Like the urbanization rate, educational attainment is recorded in the Decennial Census, meaning that missing values must be estimated by linear interpolation.

Gross domestic product

The gross domestic product (GDP) of each state is recorded annually by the Bureau of Economic Analysis, which is a government agency in the United States. Up until 1997 the bureau used the Standard Industrial Classification (SIC) system to categorize economic activity but changed to the newer North American Industry Classification System (NAICS) after 1997 (Bureau of Economic Analysis, 2023). The fact that the methodology to record this variable changed during my sample years, makes comparisons from before and after the change imperfect and is a weakness of the analysis.

Industry size of state economy

I have included three different variables to measure how the importance of different sectors in each state changes over time. These are the size of the agricultural, the manufacturing, and the service sector as a percent of the whole state economy. As with the state GDP, this data comes from the Bureau of Economic Analysis (2023) and shares the weakness that the recording methodology changed in 1997. In particular, the service sector is categorized differently in the two classification systems. I have matched the different categories that the SIC uses for the service sector as closely as possible with the new categories in the NAICS, but, as the categories do not correspond exactly, the match is not ideal.[10]

Percent of residents with low income

The US Census Bureau categorizes and documents how many people live in low-income households in the different states every year. See US Census Bureau (2023b) for details on how the Census sets the income threshold for a family to have low-income status.

Political control of the state legislature

The state legislature in all states except for Nebraska is bicameral, meaning that it is made up of two houses that both play a role in the legislative process. A political party can control (i.e., have a majority of seats) in both chambers, or the control can be split so that one party controls one chamber and another party controls the other. My data is coded such that if the Democratic Party controls both chambers, the variable takes the value of 1, if the Republican Party controls both houses it takes the value of 0, and if one party controls each camber, it takes the value of 0.5. As the Nebraska legislature is bipartisan, no value is recorded for the state in any year. The historical partisan composition of the state legislatures has been manually replicated from the National Conference of State Legislatures' annual overview of the US state legislatures (National Conference of State Legislatures, 2023).

Political leanings of the state population

[10] In the SIC, the following activities were categorized into the service sector: hotels and other lodging places, personal services, business services, automotive repair services and parking, miscellaneous repair services, motion pictures, amusement and recreation services, health services, legal services, educational services, social services, membership organizations, other services, and private households. The categories that I use from the NAICS are: professional and business services, educational services, health care and social assistance, arts, entertainment, recreation, accommodation and food services, other services (except public administration).

To get a measure for the political leanings of the state population, I use voting patterns in the national presidential elections. I chose the national election over local elections because political candidates from the same party can differ substantially in their politics based on the region where they run for office. I have taken raw vote totals from the Clerk of the United States House of Representatives' official vote count of the individual states' vote tallies published after every individual election (Clerk of the United States House of Representatives, 2022). Any votes casted for third parties (any party or candidate not from the Republican or the Democratic Party) are excluded from the analysis. The variable is then calculated for each state and year as the percent ratio of votes for the Democratic Party over the total amount of votes cast for either the Democratic or the Republican party. Thus, the variable would be 100 if all (non-third-party) voters cast their ballots for the democrats, and 0 if all votes went to the republicans. As presidential elections only give us datapoints for the partisan leanings of the states every four years, in non-election years the political preferences of the population are assumed to be an interpolation of the leanings displayed in the previous and upcoming elections.

6 Main Results

Table 2 presents the results from five different regression models. The model in column 1 is a simple OLS model of the relationship between the real minimum wage and the number of patents granted in a state. The model in column two is a TWFE DiD model with time fixed effects and state fixed effects. Column three adds demographic control variables, column 4 adds economic controls, and political control variables are also included in the model in column 5.

Table 2: Difference-in-Differences Results

Variables	(1) Log patents	(2) Log patents	(3) Log patents	(4) Log patent	(5) Log patents
Real min. wage	0.1722***	0.0904	0.0828*	0.0936**	0.0851**
	(0.0398)	(0.0557)	(0.0434)	(0.0413)	(0.0370)
Log population			1.6981***	1.4935***	1.5324***
			(0.2391)	(0.2321)	(0.242)
Urban %			0.0175	0.0185	0.0183
			(0.0147)	(0.0134)	(0.0131)
College %			0.023*	0.0232*	0.03*
			(0.0119)	(0.0125)	(0.0169)
Log GDP				0.2277	0.1712
				(0.1419)	(0.1514)
Service %				0.0109	0.0094
				(0.0159)	(0.0156)
Manufacturing %				0.0140	0.0142
				(0.013)	(0.0127)
Agriculture %				-0.0386	-0.0499
				(0.0259)	(0.0312)
Low income %				-0.0036	-0.005
				(0.006)	(0.0054)
Dem. control					0.0477
					(0.0495)
Dem. lean %					0.0017
					(0.0057)
Observations	1 650	1 650	1 650	1 650	1 617[11]
Time FE	NO	YES	YES	YES	YES
State FE	NO	YES	YES	YES	YES
Prob > | t |	0.000	0.111	0.062	0.028	0.026

Standard errors clustered by state in parentheses
Variables in percent form range from 0 to 100 percent of the state
Patents are shifted two years back to account for a delay in patent production
Real min. wage is the effective minimum wage in 2020 dollars
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

[11] The reduced sample size in model 5 is because the state of Nebraska has a bipartisan legislature and is thus dropped when the political control variable is included.

The OLS model in column 1 is to be used as a benchmark for the other, more complex, models. It shows that, when not controlling for any other variables or fixed effects, there is a positive relationship between the minimum wage and patent production. It is probable however, that the model cloaks biases that could significantly distort the results. As such, the model is too simple to draw any causal inference from.

When using a DiD model, the effect is reduced, but still positive – though the results are now not statistically significant at the $p < 0.1$ level. This is shown in column 2 of table 2. This model shows that when taking into account the differences that were already present before any changes in the minimum wage and time trends that affected all states (which are captured here by the fixed effects), the effect is weakened. Nevertheless, bias in the model could still be present because, while this form of DiD does adjust for all omitted variables affecting the coefficient *before* any changes in the wage, we do not yet consider the effects that differences in omitted variable's trends *during* the observed period could have.

The model in column 3 adds control variables in panel data form for three demographic indicators: population (log form), the urbanization rate, and the percent of the population with a college degree or higher in each state. By adding these variables, the model controls for how the trends of these covariates impacted patent production in the observed time period. Doing so makes the effect of minimum wages on innovation significant at the $p < 0.1$ level. A correlation coefficient of 0.0936 between the real minimum wage and the log transformed number of patents can be interpreted as that raising the minimum wage by 1 USD would cause patent production to increase by around 9 percent. The current federal minimum wage is at 7.25 USD per hour, so an increase by the size of 1 USD would be a significant change. Unsurprisingly, we see that there is a positive relationship between the size of the population and the number of patents granted – as the population grows, so does the number of patents. The fact that this number is above one (1.4935), implies that there are "scale benefits" to having a large population in terms of the impact on patents, or at the very least that the states who grew more were also the states whose patent production grew more in proportion to their capita. The parameter estimate of the education indicator is also positive and significant. This indicates a positive relationship between the general education level in the population and patent production. In contrast, the measure of urbanization is not significant at the 0.1 level.

Model 4 and model 5 are similar to model 3 but controlling for additional trends. Model 4 adds the economic variables of GDP (log form), the fraction of the population with a low income, and what percent of the state economy is in the service sector, the manufacturing sector, and the agricultural sector. The model in column 5 also includes two political indicators: what political party controls the state legislature and how the citizens of the state vote in general elections. The two models give similar results. The coefficient of the minimum wage is in both cases around 0.09, and the p-value is around 0.03. None of the control variables added by these two models have a statistically significant effect on the dependent variable on their own. Following evidence from Altonji et al. (2005), the results are likely to remain largely unchanged if additional control variables were to be added owing to the fact that the observable covariates modestly affect the outcome.

7 Statistical Validity

In this chapter, I will first investigate the parallel trends assumption and whether it seems to be fulfilled for the analysis of my thesis. Second, treatment effect heterogeneity will be explored, and alternative estimators to the TWFE estimator will be examined in order to explore whether heterogeneous treatment effects invalidate the results.

7.1 Parallel Trends

As previously outlined in chapter 4, the parallel trends assumption is an essential requirement for a DiD estimation. The assumption is that the control group shows what would have happened to the dependent variable of the treatment group if the treatment group had the same level of treatment as the control group. One way to examine the likelihood of this assumption is to look at whether the two groups display parallel trends, in this case in innovation, in the past. This can be done with an event study. An event study (in the context of this paper) shows how a control group and a treatment group differ over time – before and after an event (in this case being the raising of the minimum wage). Ideally the difference between the two should be close to zero in the periods before the event for the parallel trends assumption to hold.

An event study cannot *prove* that the parallel trends assumption is satisfied, as the fact that trends have been parallel in the past does not necessarily *guarantee* that trends in the future would also have been parallel, but it does provide suggestive evidence that the assumption is unlikely to be violated. On the other hand, if the result shows that the two groups' trend have not been parallel in the past, this is a warning sign that the control group is an inadequate counterfactual for the treatment group's trend.

As event studies requires there to be "events", i.e, a binary change in the treatment, the estimator used for the event study in figure 2 is the Callaway and Sant'Anna estimator described in chapter 7.2.1.

Figure 2: Event Study on Minimum Wage and Patents

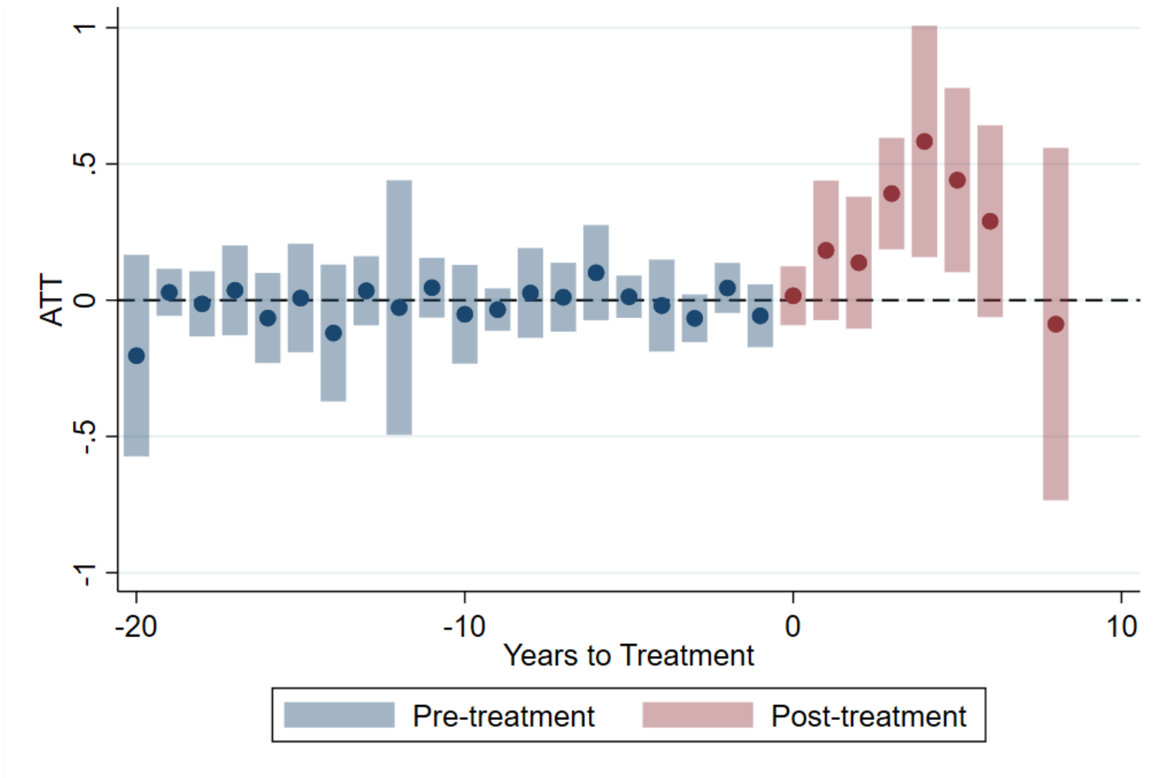

Figure 2 compares the states before and after a minimum wage is raised above the federal minimum. As we can see, the difference appears to be close to zero in the years before the change, while there is a noticeable increase in patent production after the minimum wage is raised. More specifically, the *ATT* has a coefficient of 0.18. One important thing about the graph, which I will discuss more fully in chapter 7.2.1, is that one should be careful when interpreting the magnitude of the effect of the Callaway and Sant'Anna estimator in the post-treatment period, as the "dosage level" of minimum wage is ignored in this specification, and instead recategorized as a dummy. The post-treatment period therefore shows the average effect of raising the minimum wage, regardless of how sharp the increase was. The graph is therefore not intended to be used to gauge the magnitude of the effect – one should instead focus on the pre-treatment portion of the graph. The latter is unaffected by the discretization of the independent variable, as the treatment level is equal for all states in this period regardless. The fact that the treatment effect in this part of the graph appears to be nonexistent, shows us that control and treatment groups have parallel trends before there are any differences in treatment, and makes it more likely that the trends would have continued to be parallel if not for the treatment.

It should be noted that, while the event study in figure 2 can be used to argue for the validity of the traditional parallel trends assumption, one should be more careful in one's claims of what the graph can demonstrate about "strong" parallel trends. See chapter 4.2 for a discussion on the credibility of strong parallel trends in this setting.

7.2 Treatment Effect Heterogeneity

As discussed in chapter 4.2.1, when the treatment occurs at differing times between the units, the TWFE estimator might be biased if there are also heterogeneous treatment effects. If this is the case, and the treatment effect varies across units or over time, the regression estimates weighted sums of the treatment effect which may be negative and cause misleading results (de Chaisemartin & D'Haultfœuille, 2020).

In my analysis so far, the treatment timing of when the states raise their minimum wages is variable. This means that the main results of this paper could suffer from negative weights and be biased. To account for this problem, I will explore alternative estimators to the TWFE estimator which allows for treatment effect heterogeneity and see how these affect the results of the regression.

7.2.1 Alternative Estimators

In the paper *Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects*, Clément de Chaisemartin and Xavier D'Haultfœuille explore the limitations of the commonly used TWFE estimator in panel data settings with heterogeneous treatment effects. The authors argue that the TWFE estimator may produce biased estimates under the circumstances mentioned above, and they propose an alternative approach to address this issue. Their estimator uses only those observations where there exists both what they call a "joiner" or a "leaver" – i.e., a unit going from one treatment level to another – and a unit which is at the original treatment level in both periods. The estimator then compares the evolution of the mean outcome between the groups in periods t – 1 and t. With a continuous treatment one can also specify a threshold for how big of a difference to allow in the treatment variable before it is classified as being on another treatment level.

When using the Chaisemartin and D'Haultfœuille estimator instead of the TWFE estimator, the coefficient of the independent variable remains positive (0.0430). However the results are no longer statistically significant. This is likely because the sample size is dramatically reduced when using this method (the sample size with the Chaisemartin and D'Haultfœuille estimator is 94, as opposed to 1617 in model 5). Because of the small sample size, this method fails to shed much light on how the estimator might be biased by heterogeneous treatment effects. See appendix A.1 for more details on the results when using this estimator on the data.

Another estimator that is robust to heterogeneous effects is the one presented in the paper *Difference-in-Differences with Multiple Time Periods* by Brantly Callaway and Pedro H. C. Sant'Anna (2021). This estimator classifies units into groups of "never-treated", "always-treated", "not-yet-treated", and "already-treated" in every time period. Always-treated units are eliminated entirely from the analysis, while already-treated units are excluded from the control group after they have been treated. Using only the comparisons that are left after excluding the inappropriate controls, the estimator finds the group-time ATT for all cohorts treated at the same point in time.

Using the estimator from Callaway and Sant'Anna requires two modifications to my main analysis using the TWFE estimator. First, the treatment must be binary. Second, the treatment must be irreversible, meaning that units cannot become untreated again after having been treated before. Reclassifying the treatment variable (the minimum wage rate) into a binary dummy variable is not without problems. As previously discussed, all states in the US must at least have the federal minimum wage floor, with the possibility of a higher local minimum wage. This means that all states are treated to some extent, and so using a binary treatment can at best show the effect of having a wage higher than the federal minimum, not the effect of having a minimum wage at all. The magnitude of the effect also becomes difficult to interpret when the treatment is reclassified as being zero when at the federal level and one when it is higher. Grouping all those who have a minimum wage "above the federal minimum" masks the variation in effective minimum wages across states. Hence, it becomes somewhat unclear exactly what level of minimum wage it is that causes the effect we observe. Nevertheless, while the magnitude of the effect might be harder to interpret with this estimator than with a TWFE continuous estimator, it is still useful to explore whether the direction of the effect is the same when we use an estimator which is robust to heterogeneous effects.

The Callaway and Sant'Anna estimator measure the ATT using the following formula:

$$ATT(g,t) = \mathbb{E}\left[\left(\frac{G_g}{\mathbb{E}[G_g]} - \frac{\frac{p_{g,t}(X)(1-D_t)}{1-p_{g,t}(X)}}{\mathbb{E}\left[\frac{p_{g,t}(X)(1-D_t)}{1-p_{g,t}(X)}\right]}\right)\left(Y_t - Y_{g-1} - \mathbb{E}[Y_t - Y_{g-1}|X, D_t = 0, G_g = 0]\right)\right]$$

The subscripts g and t indicate group and time. The dosage D is in this estimator binary. G indicates in what period the unit is first treated, and the propensity score p is the probability of receiving treatment, given the measured covariates X.

Table 3 shows the result from the DiD regression on my data when using Callaway and Sant'Anna's estimator and a binary treatment. Because the treatment must be irreversible in Callaway and Sant'Anna's DiD design, I have dropped states from the dataset after they have become untreated, i.e., after the federal minimum wage has "caught up" with the state minimum wage.

Table 3: Results from the Callaway and Sant'Anna (2021) DiD Estimator

	Coefficient	Std. err.	z	P > | z |	95% conf. interval	
ATT	0.1819547	0.0888402	2.05	0.041	0.007831	0.3560784

Outcome model: least squares; treatment model: inverse probability
Standard errors clustered by state
Number of observations: 1 146

The coefficient of the ATT is positive, and the regression is statistically significant at the $p < 0.05$ level. This indicates that there is still an effect of minimum wages on patents when taking into consideration that treatment effects can be heterogeneous.

If one were to try to give a broad interpretation of the magnitude of the coefficient, it is necessary to unpack the actual size of the minimum wage, which in the model is coded as a dummy variable. Averaging the treatment states, the minimum wage increased by 0.83 USD – from an average rate of 7.58 USD to an average of 8.41 USD following treatment. In the same period the average minimum wage rate had a slight decrease of 0.07 USD for the control group – from 7.58 USD to 7.51 USD. It appears like, when we take heterogeneous treatment effects into account, the difference of 0.9 USD between the treatment and the control group caused the treatment group to produce around 18 percent more patents than the control group, which is a stronger effect than what was found when heterogenous treatment effects were not considered. Keep in mind though that, due to the treatment being discretized, variation which could give valuable information is lost, and the estimator could be less accurate as a result.

8 Discussion

The empirical analysis shows a positive relationship between a wage floor and innovation. In the theoretical framework of Acemoğlu (2010), this implies that technology is labor-saving. As explained in section 2.2, technology can be labor-saving or labor-complementary depending on the returns to scale, the elasticity of substitution between labor and capital, and the degree to which technology is labor-augmenting and capital-augmenting. Knoblach et al.'s (2019) meta-nalysis show that the long-run elasticity of substitution is consistently estimated to be between 0.45 and 0.87 in the United States. Given my empirical results, this suggest that technology largely improves the productivity of labor and that the returns to scale are sufficiently diminishing for increased labor costs to incentivize more technology use.

However, this analysis only concerns the macroeconomic level. On the scale of individual industries – and possibly also of individual states – these conditions are likely to differ, leading to a more diffuse relationship between wage floors and innovation. Both the returns to scale and the elasticity of factor substitution are dependent on the specifics of the market, and there could therefore be heterogeneity in the responsiveness of innovation to increased labor costs.

To evaluate to what extent such heterogeneity is present, consider Appendix A.2. It contains the results of a DiD estimation on the minimum wage increases of 2007, which was a year when several states raised their minimum wage at the same time. When looking at this case in isolation, the statistical uncertainty is too high to determine the effect of the minimum wage on patents. It is possible that differing characteristics of this specific event with respect to the general economic environment (where we observe an effect) weaken, or even revert the relationship, so that I am unable to find the same effect. The year 2007 was right before the great recession, and so it is conceivable that the higher labor costs were too much for firms to bear in the harsh economic environment and hurt states with a high minimum wage more. Though there seems to be a general positive effect of wage costs on innovation, I cannot conclude that there are no time and group differences in the effect. Hence, it is not guaranteed that a high minimum wage will *always* encourage innovation.

Additionally, the empirical analysis is done on US data, and as such the results may not extrapolate to other settings. As explained in chapter 3, one important difference between the United States and other countries is that elsewhere it is common to establish wage floors through collective bargaining as opposed to through a statutory minimum wage. This has consequences on how wage floors affect innovation. Besides the fact that union wages can be more sensitive to the individual industry, one important distinction is how unions affect high- and low-wage jobs differently. Historically, unions have shrunk the income gap between the top and bottom earners, in practice holding back the wage of high-earners while increasing the wage of low-earners (Barth et al., 2014). To the extent that technological-complimentary labor earns high wages (supported by the capital-skill complementarity hypothesis (Krusell et al., 2000)) and that low-wage labor is more easily replaceable by technology, unions will probably encourage innovation to a higher degree than a minimum wage will.

9 Conclusion

The empirical analysis conducted in this master's thesis has shown a positive effect of minimum wages on innovation, as reflected in patent production. The TWFE estimator shows that a 1 USD increase in the minimum wage leads to approximately a 9 percent increase in patent production, though this effect could be larger when accounting for heterogeneity in treatment effects, as well as possible spillover effects from treatment states to states with a low minimum wage. The result is congruent with Habakkuk’s theory that labor scarcity encourages technological innovation.

However, to a certain degree, these findings are likely context-specific. Replicating the analysis in other settings may lead to different results if the economic conditions and the institutional background differ considerable to the one in the United States. It is crucial to take into account differences in the responsiveness of technological innovation due to the elasticity of substitution between labor and capital and the rate of returns to scale. These nuances could alter the observed effect, with the potential for both larger and smaller impacts of minimum wage changes on patent production.

Despite these considerations, the overall positive effect identified in this study lends support to Habakkuk's view, which underscores the potentially crucial role of labor costs in driving innovation. This establishes a noteworthy link between wage floors and technological advancements that warrant further examination.

As for future research, it would be instructive to further explore the conditions that shape the relationship between wage floors and innovation. In particular, examining markets beyond the United States would provide a more global perspective on this matter. Furthermore, it would be interesting to explore how unions, another factor influencing labor costs, compare to minimum wages in terms of their impact on innovation. As technological progress is a vital driving force of economic growth, its relationship with labor costs is an area of research deserving of attention.

References

Abbott, D. W., & Levine, J. P. (1991). *Wrong Winner: The Coming Debacle in the Electoral College*. Praeger.

Acemoğlu, D. & Finkelstein, A. (2008). Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector. *Journal of Political Economy, 116*(5), 837–80. https://doi.org/10.1086/595014

Acemoğlu, D. (2010). When Does Labor Scarcity Encourage Innovation? *Journal of Political Economy, 118*(6), 1037-1078. https://doi.org/10.1086/658160

Acemoğlu, D., & Autor, D. (2011). Skills, Tasks and Technologies: Implications for Employment and Earnings. *Handbook of Labor Economics*, *4*, 1043–1171. https://doi.org/10.1016/s0169-7218(11)02410-5

Acemoğlu, D. & Restrepo, P. (2020). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy, 128*(6). https://doi.org/10.1086/705716

Allen, R. C. (2009). *The British Industrial Revolution in Global Perspective*. Cambridge University Press.

Altonji, J. G., Elder, T. E., & Taber, C. R. (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy*, *113*(1), 151–184. https://doi.org/10.1086/426036

Arpaia, A., Cardoso, P., Kiss, A., Van Herck, K. & Vandeplas A. (2017). Statutory Minimum Wages in the EU: Institutional Settings and Macroeconomic Implications. *IZA Policy Paper No. 124 from Institute of Labor Economics*. https://docs.iza.org/pp124.pdf

Barth, E., Moene, K. O., & Willumsen, F. (2014). The Scandinavian model—an interpretation. *Journal of Public Economics, 117*, 60–72. https://doi.org/10.1016/j.jpubeco.2014.04.001

Bureau of Economic Analysis (2023) *Regional Data: GDP and Personal Income*. United States Department of Commerce. https://apps.bea.gov/itable/?ReqID=70&step=1#eyJhcHBpZCI6NzAsInN0ZXBzIjpbMSwyNF0sImRhdGEiOlt-bIlRhYmxlSWQiLCI1MDUiXV19

Bureau of Labor Statistics. (2023a). *Consumer Price Index (CPI) Databases*. United States Department of Labor. https://www.bls.gov/cpi/data.htm

Bureau of Labor Statistics. (2023b). *Union Members – 2022*. United States Department of Labor. https://www.bls.gov/news.release/pdf/union2.pdf

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics, 225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Callaway, B., Goodman-Bacon, A. & Sant'Anna, P. H. C. (2021). Difference-in-Differences with a Continuous Treatment. *arXiv:2107.02637*. https://doi.org/10.48550/arXiv.2107.02637

Chin, J. C. (1998). Rural-Urban Wage Differentials, Unemployment, and Efficiency Wages: An Open Economy Policy Analysis. *Southern Economic Journal, 65*(2), 294–307. https://doi.org/10.2307/1060669

Clemens, M. A., Lewis, E. G., & Postel, H. M. (2018). Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion. *American Economic Review, 108*(6), 1468– 1487. https://doi.org/10.1257/aer.20170765

Clerk of the United States House of Representatives (2022). *Election Statistics: 1920 to Present*. US House of Representatives: History, Art & Archives. https://history.house.gov/Institution/Election-Statistics/

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/10.1257/aer.20181169

Dechezleprêtre, A., Hémous, D., Olsen, M. & Zanella, C. (2019). Automating Labor: Evidence From Firm-Level Patent Data. *SSRN*. http://dx.doi.org/10.2139/ssrn.3508783

de Rassenfosse, G. (2013). Do Firms Face a Trade-Off Between the Quantity and the Quality of their Inventions? Research Policy, 42(5), 1072–1079. https://doi.org/10.2139/ssrn.2228993

Desilver, D. (2021, May 20). The U.S. Differs from Most Other Countries in How it Sets its Minimum Wage. *Pew Research Center*. https://www.pewresearch.org/short-reads/2021/05/20/the-u-s-differs-from-most-other-countries-in-how-it-sets-its-minimum-wage/

Ekeland, A., Pejarinen, M. & Rouvinen P. (2015). *Computerization Threatens One-Third of Finnish and Norwegian Employment*. (ETLA Brief No. 34) The Research Institute of the Finnish Economy. https://www.etla.fi/en/publications/computerization-threatens-one-third-of-finnish-and-norwegian-employment/

Elvin, M. (1972). *The High Level Equilibrium Trap: The Causes of the Decline in Invention in the Traditional Chinese Textile Industries*. Stanford University Press.

Fair Labor Standards Act of 1938 (1938). https://www.dol.gov/sites/dolgov/files/WHD/legacy/files/FairLaborStandAct.pdf

Figueroa, A. (2022, June 6). Advocates for $15-an-hour Federal Minimum Wage Press Biden for a Meeting. *New Jersey Monitor*. https://newjerseymonitor.com/2022/06/06/advocates-for-15-an-hour-federal-minimum-wage-press-biden-for-a-meeting/

Fritsch, M., & Wyrwich, M. (2021). Is Innovation (Increasingly) Concentrated in Large Cities? an international comparison. *Research Policy, 50*(6). https://doi.org/10.1016/j.respol.2021.104237

Fölster, S. (2018). *Norway's New Jobs in the Wake of the Digital Revolution*. NHO.

https://www.nho.no/publikasjoner/arbeidsliv/norways-new-jobs-in-the-wake-of-the-digital-revolution/

Habakkuk, H. J. (1962). *American and British Technology in the Nineteenth Century: The Search for Labor-Saving Inventions*. Cambridge University Press.

Hannan, T., & McDowell, J. (1984). The Determinants of Technology Adoption: The Case of the Banking Firm. *RAND Journal of Economics, 15*(3), 328–335.

Hicks, J. (1932). *The Theory of Wages*. Macmillan.

James, J. A. & Skinner, J. S. (1985). The Resolution of the Labor-Scarcity Paradox. *The Journal of Economic History, 45*(3), 513-540. https://www.jstor.org/stable/2121750

Knoblach, M., Roessler, M., & Zwerschke, P. (2019). The Elasticity of Substitution Between Capital and Labour in the US Economy: A meta-regression analysis. *Oxford Bulletin of Economics and Statistics*, *82*(1), 62–82. https://doi.org/10.1111/obes.12312

Krusell, P., Ohanian, L. E., Rios-Rull, J.-V., & Violante, G. L. (2000). Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis. *Econometrica*, *68*(5), 1029–1053. https://doi.org/10.1111/1468-0262.00150

Lordan, G. & Naumark, D. (2018). People Versus Machines: The Impact of Minimum Wages on Automatable Jobs. *Labour Economics, 52*, 40-53. https://doi.org/10.1016/j.labeco.2018.03.006

Manuelli, R. E., & Seshadri, A. (2014). Frictionless Technology Diffusion: The Case of Tractors. *American Economic Review, 104*(4), 1368–1391. https://doi.org/10.1257/aer.104.4.1368

National Conference of State Legislatures. (2023, May 23). *State Partisan Composition*. NCSL. https://www.ncsl.org/about-state-legislatures/state-partisan-composition

National Agricultural Law Center. (2022, January 25). *Minimum Wage for Agricultural Work-*

ers. University of Arkansas System Division of Agriculture. https://nationalaglaw-center.org/state-compilations/agpay/minimumwage/

Nain, A. & Wang, Y. (2021). The Effect of Labor Cost on Labor-Saving Innovation. *SSRN*. http://dx.doi.org/10.2139/ssrn.3946568

Neumark, D. (2017). The Employment Effects of Minimum Wages: Some Questions We Need to Answer (No. w23584). *National Bureau of Economic Research*.

Ricardo, D. (1951). *Works and Correspondences.* Edited by P. Sraffa. Cambridge University Press.

Rothbarth, E. (1946). Causes of the Superior Efficiency of U.S.A. Industry as Compared with British Industry. *The Economic Journal, 56*(223), 383-390. https://doi.org/10.2307/2226046

Slinn, S. J. (2023). Workers' Boards: Sectoral Bargaining and Standard-Setting Mechanisms for the New Gilded Age. *Employee Rights and Employment Policy Journal, Forthcoming*. https://ssrn.com/abstract=4429598

Statista Research Department (2023, May 3). *Number of workers paid hourly rates with earnings at or below the minimum wage in the U.S. in 2021, by industry*. Statista https://www.statista.com/statistics/299372/us-minimum-wage-workers-by-industry/

Temin, P. (1966). Labor Scarcity and the Problem of American Industrial Efficiency in the 1850's. *The Journal of Economic History, 26*(3), 277-298. https://www.jstor.org/stable/2115648

United States Department of Labor. (2023, January 1). *Changes in basic minimum wages in non-farm employment under State Law: Selected Years 1968 to 2022.* DOL. https://www.dol.gov/agencies/whd/state/minimum-wage/history

United States Patent and Trademark Office (2015). *Extended Year Set - Patent Counts By*

Country, State, and Year: All Patent Types. https://www.uspto.gov/web/offices/ac/ido/oeip/taf/cst_allh.htm

United States Patent and Trademark Office (2023, May 2). *Patent Basics*. USPTO. https://www.uspto.gov/patents/basics

US Census Bureau (2023a). *Census Data*. https://data.census.gov/

US Census Bureau. (2023b, January 30). *How the Census Bureau Measures Poverty*. Cen sus. https://www.census.gov/topics/income-poverty/poverty/guidance/poverty-measures.html

World Intellectual Property Organization (2022). *Global Innovation Index 2022: What is the Future of Innovation-driven Growth?* WIPO. https://doi.org/10.34667/tind.46596

Appendix

A.1 The Chaisemartin and D'Haultfœuille Estimator

The Chaisemartin and D'Haultfœuille Estimator is robust against treatment effect heterogeneity. In order to be able to compare units with each other, I recategorized the treatment (minimum wage) to group the states by one dollar wide treatment chunks. The threshold of "allowable treatment variation" was set to 0.4.

Table 4: Results from the Chaisemartin and D'Haultfœuille (2020) DiD Estimator

	Coefficient	Std. err.	z	P > | z |	95% conf. interval	
ATT	0.0430407	0.0367695	1.17	0.2418	-0.029028	0.1151089

Standard errors clustered by state
Number of observations: 94

The results show a positive coefficient of the treatment to the outcome, but the results are not statistically significant, likely due to the dramatically reduced sample size.

A.2 Case Study: The Minimum Wage Increases in 2007

To investigate the minimum wage increases of 2007, I have included only those states that either 1) first raised their minimum wage in 2007 and kept it above the federal level for at least eight years or 2) kept their minimum wage at the federal level until at least eight years after 2007. The former functions as the treatment group while the latter is the control group. The pre-treatment period is limited to eight years and the same is true the post-treatment period. The choice of limiting the study to eight years before and after the change, means that we can include Nebraska and South Dakota in the sample, who both raised their minimum wage in 2015, though Montana must be excluded from the sample, as they increased their minimum wage in 2014. This specification ensures that the estimator remains unbiased from the problematic aspects of DiD designs with staggered treatment.

The states in the treatment group are Arizona, Colorado, Michigan, and Ohio.

The states in the control group are Alabama, Georgia, Idaho, Indiana, Kansas, Kentucky, Louisiana, Mississippi, Nebraska, North Dakota, Oklahoma, South Carolina, South Dakota, Tennessee, Texas, Utah, Virginia, and Wyoming.

To check for parallel trends, I use graphical diagnostic plots. Two plots are presented in figure 3. The initial graph includes a pair of trend lines that display the average outcome over time for both the treatment and the control groups. The subsequent graph extends the DiD model to incorporate time interactions with a treatment indicator and plots the projected values from this augmented model for the treatment and control groups. Table 5 shows the results of the estimator.

Figure 3: Graphical Diagnostics for Parallel Trends

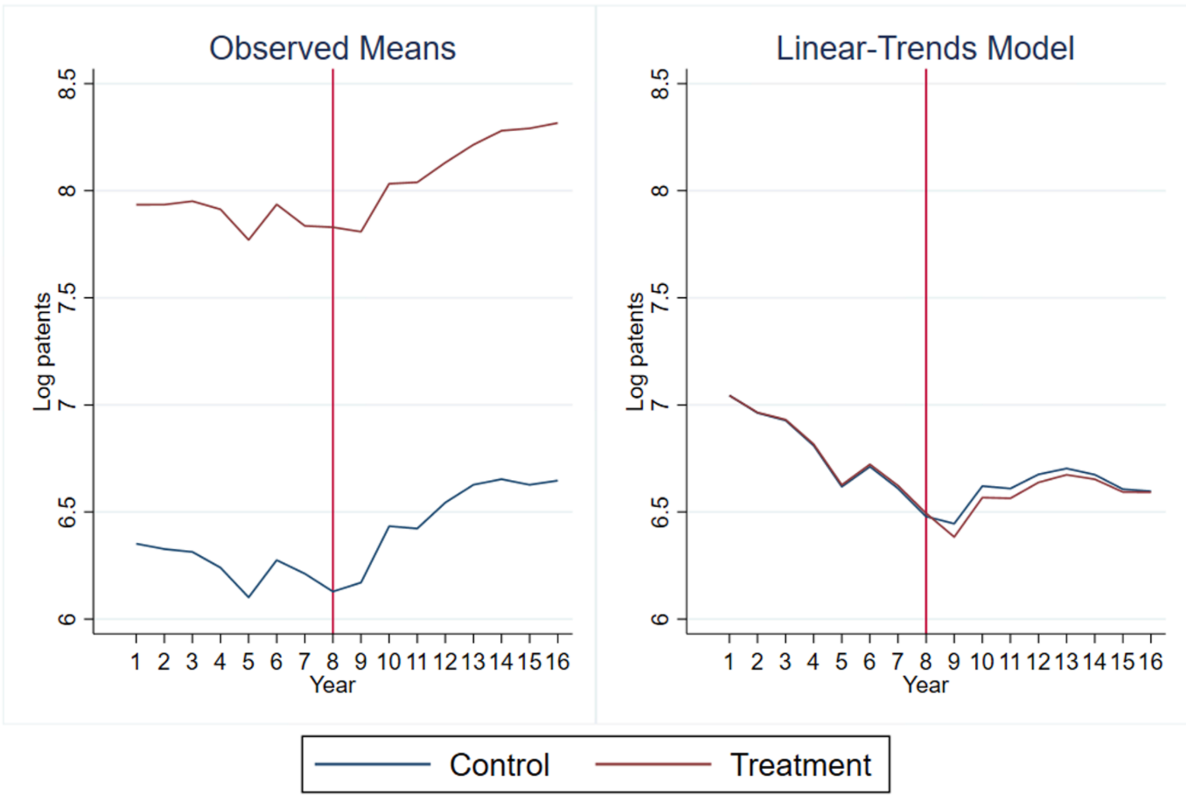

Table 5: Results from the TWFE Estimator on the 2007 Sample

	Coefficient	Std. err.	t	P > | t |	95% conf. interval	
ATT	-0.011614	0.0243132	-0.478	0.638	-0.06233	0.0391028

Standard errors clustered by state
Number of observations: 336

The coefficient of the treatment is negative, and the results are not statistically significant.