Berkeley Economic Review

| Content

Professor Interviews

Research Papers

| Peer Review Board

SUBMISSIONS POLICY

- Format: Please format your submission as a Microsoft word document or Latex file. We would prefer submissions single-spaced in size 12, Times New Roman font.
- Length: Maximum five pages single-spaced for essays or op-eds, not including works cited page(s). No maximum for theses or original research papers.
- Plagiarism: We maintain a strict zero tolerance policy on plagiarism. According to the University of North Carolina, plagiarism is "the deliberate or reckless representation of another's words, thoughts, or ideas as one's own without attribution in connection with submission of academic work, whether graded or otherwise." For more information about plagiarism—what it is, its consequences, and how to avoid it—please see the UNC's website.
- Citations: Use MLA format for in-text citations and your works cited page. Include a separate works cited page at the end of your work.
- Send to: berkeleyeconreview@gmail.com

DISCLAIMER

The views published in this journal are those of the individual writers or speakers and do not necessarily reflect the position or policy of Berkeley Economic Review or the University of California at Berkeley.

COPYRIGHT POLICY

STAFF

Editors in Chief

Vatsal Bajaj*
Vinay Maruri*

Peer Review

Charlie McMurry
Christian Philip Hoeck
Danielle Chen
Emaan Siddique
Elzette Janse van Rensburg
Hayden Davila
Jules Baudet
Patrick Tagari*
Renuka Garg
Taylor Wang
Xilin Ying

Research and Editorial

Qualitative Staff Writers

Abhishek Roy
Konnor von Emster
Olivia Gingold
Matt Arena
Nanditha Nair
Pak Hei So
Raina Zhao
Yash Rajwanshi
Yizhuo Shen

STAFF

Quantitative Staff Writers

Andreas Maass
Grace Jang
Jeffrey Suzuki
Vanessa Thompson

Editors

Amanda Yao
Alex Cheng
Andrew Babson
David Lyu
Gillian Sero Del Mundo
Lucia Dardis
Howard yan
Odysseus Pyrinis
Parmita Das*
Shawn Shin*

Layout and Design

Anne Fogarty*
Jaide Lin
Natalia Ramirez
Sasha Hassan
Shengning Zhang

Web Design

Gurmehar Somal*
Prakash Srivastava
Richard Gong

STAFF

External Affairs

Selena Zhang*

Outreach

Ally Mintzer
Camila Reimpell*
Douglas Koehler
Hugo Letouze
Niolas Dussaux
Pearleen Wang
Sahil Dang

Marketing

Britney Nicolas
Evelina Bergstrom*
Heesung Park

Professional Development

Jiangdong Fang*
Lizbeth Flores
Sophia Li

Additional Staff

Finance
Rishi Modi*

Social Chair
Mae Zhang*

Historian
Shunsuke Kinoshita

** Officers*

| From The Editors' Desk

Dear BER Reader,

On behalf of the 65 members in our Research and Editorial, Peer Review, Layout & Design, External Affairs, Web Development, and Executive Teams, we are proud to present the eighth edition of the Berkeley Economic Review.

As the voice of economics undergraduates around the world, the Berkeley Economic Review presents five professor interviews and four distinguished undergraduate research papers which tackle economic dilemmas facing the modern world.

First, we present Saleel Huprikar's investigation on whether a person's geographical location impacts their perspectives on the federal minimum wage. Second, Brian O'Connor discerns the impact of the UK's National Living Wage on employment probabilities of low wage workers. Huprikar and Connor's research advances knowledge about one of the most hotly debated topics in labor economics, minimum wage and its impact on the distribution of income.

Third, Laure Fleury analyzes the impact of the knowledge economy factors on total factor productivity. As economies around the world stagnate, Fleury's research investigates the reasons behind residual economic growth. Finally, Tao Chen examines the impact of first degree price discrimination and quality customization under data protection regulations. In a world with growing privacy concerns due to artificial intelligence, big data, and mass surveillance, Chen investigates the aims and effectiveness of data protection regulations.

We hope that you will immerse yourself in the Berkeley Economic Review and gain new perspectives on economic solutions to global issues. In this academic spirit, we present to you the eighth edition of Berkeley Economic Review.

Best,
Vatsal Bajaj and Vinay Maruri
Editors-In-Chief, *Berkeley Economic Review*

Professor Clair Brown

Interviewed by ***Jeffrey Suzuki***

Interviewer: Hi professor! If you don't mind, please introduce yourself to our readers. Why did you choose economics?

Prof. Brown: I saw economics as a field where people could really make a difference. The Women's movement and the Civil Rights movement were both in full swing then. It was a field where people could change the world. For my dissertation, I studied how discrimination functioned in labor markets. Economics is an area where people could measure injustices in living standards and inequality. And this was all before we knew that there was a climate problem! And now we have economists studying this issue as well. Right now, I have a team of researchers studying sustainable policies across the world!

Interviewer: You are known for being an outspoken proponent of what you call "Buddhist Economics." What exactly is "Buddhist Economics?"

Prof. Brown: Buddhist economics is built upon three ideas; the first two are ecological assumptions. First, humans have interdependence with others and the planet—a buddhist idea—but also one that scientists have supported. It could be understood as the 4th Law of Ecology. Secondly, Buddhist Economics argues impermanence: the idea that nothing lasts forever. This is also a view espoused by scientists. Lastly, people actually care for each other in an altruistic way. This is not exactly a tenet of Buddhism as much as it has been shown by neuroscientists. All in all, Buddhist economics says that there is more to life than maximizing productivity or consumption. Because we are all interdependence and naturally care for each other, we should seek to improve the lives of others and act in accordance to our altruistic nature.

Interviewer: What would you say to a wealthy businessman who claims that he earned his way to the top, and questions why his taxes should go toward some poor kid's education? What if truly believes that he and his competitors' greed is the grease that keeps our economy's wheels spinning?

Prof. Brown: I would give two responses:

First, I would argue that he was able to do what he did because of his own education, which is a social program. His success is dependent on society's ability to educate people like himself. Additionally, our infrastructure benefits him. Society ensures through health programs that his workers are able to provide labor. Society ensures that his workers are provided an education so that they can be more effective. The government ensures that he is able to safely work in a global economy and use its technology. His productivity—his success—is dependent on the system. And all of these programs are only possible through the taxes he and everyone else pays. The idea that he did it by himself is incorrect.

Secondly, economists argue for incentives. Every time someone has done a study for the returns to effort, it's actually very little. As it turns out, people still work very hard when their taxes increase. Even if the returns to their labor decrease from taxes, they still make an effort to work hard. There are other factors that make people happy besides income, such as the general experience of working the job. Economists still argue whether we should increase progressive taxes to 72% or 76%. However, all of that aside, we had taxes like that in our history and there didn't seem to be a substantial impact. We might not want to raise taxes to 98%—we want to make sure we have the revenues to make our social programs work. To put it simply, the businessman in your example would probably still work all the same.

Interviewer: GDP has become the measure many economists use to calculate economic prosperity in the macro sense. Why do you think that this is a flawed measurement of welfare?

Prof. Brown: Because all it measures is what goes through the marketplace. Every economist will tell you it's flawed. Even Simon Kuznets, the person who developed GDP, stated that it should not be used to indicate welfare. GDP leaves everything that is currently left out of the marketplace like leisure, home improvement, child care, environmental degradation, and a bunch of other important factors. Finally, any [negative] externalities of the marketplace are left out of GDP; we are leaving

out the negatives of how we actually produce and consume. With GDP, we only measure our production and consumption.

Interviewer: Besides GDP, what other metrics are currently available for economists to use that could be superior to GDP for measuring welfare?

Prof. Brown: There are lots of metrics. That's part of the problem. It's like the saying where if you ask five economists the same question, you get five different answers. But I can give some examples. One is the Genuine Progress Indicator, or GPI, which places a market value for all activities outside of the scope of the market. So, it gathers data on things like non-market labor, the use of time, environmental degradation, and all these other aspects that are not captured by GDP and places a market value on them. However, it takes a lot of data. In fact, GDP has the same problem but we've computed it for such a long time, spending billions of dollars to ensure that this metric is gathered. But discussions around welfare usually revolve around GDP per capita, which leads many people astray.

Interviewer: How does GDP per capita lead us astray?

Prof. Brown: GDP per capita might be growing over time but it doesn't account for the distribution of income. For most of the 20th century, GDP per capita and median household incomes are going up at the same time. But there is this big divide that happened in the late 70s because Reagan began to slash taxes. Ever since, a large amount of the growth in GDP has been captured by the top 5% and 1%. Industries became more concentrated. More income and power in industries allows wealthy individuals and corporations to buy elections. The impact on the environment is ignored. It has become a vicious cycle. This is all characteristic of living in a neoliberal world.

Interviewer: Speaking of alternative measures of welfare, you have co-written an article on Project Syndicate with another researcher where you describe work on a policy index that compares 50 different countries around the world, called the Shared Sustainability and Prosperity Index (SSPI). Can you

elaborate on what exactly it is—what it measures?

Prof. Brown: I have terrific team working on it right now! We're going to expand the number of countries it covers and improve its indicators. It is still a work in progress, however. When I was having conversations about the impacts of policies, we would talk about certain countries having certain policies, such as increasing the education level of girls. We knew that certain policies had positive or negative effects. However, on a global level, I was curious about which countries actually have these policies. Which countries have policies that combat climate change? Which countries have policies that are spreading education in girls and women? Performance aside, we wanted to answer this question: "how do governments from around the world compare in terms of policy?" We try to answer this question by observing policies that indicate how countries govern, structure markets, and foster sustainability. I love markets, by the way. But it matters how we regulate and set up social safety nets to ensure that everyone can meet their potential.

Interviewer: Once the [SSPI] is finished and ready, how do you expect it to be used?

Prof. Brown: That's a great question! One of the things we learned is that these policies don't change rapidly. There are many metrics for performance metrics. We asked that do policies track well with these performance metrics. In other words, do policies matter? And we were happy to find that they do. The policies that promote responsible governance, market structure, and sustainability track very well with economic performance. Some of my colleagues have noted that this seems to pave the way for justifying the Green New Deal.

Interviewer: We're starting to hit the end of this interview, so I'll ask one last question: as you know, the UN estimates that there are only about 11 years left before emissions inflict irreparable damage to the Earth's climate. Do you believe that societies around the world will be able to mobilize before the disastrous effects of climate change become irreversible, or that humanity will have to pay a hefty cost before learning its

lesson?

Prof. Brown: I try really hard to be optimistic. However, I feel somewhat pessimistic. When people ask what we can do to combat climate change, I say that we need to fight. I worked with UC Regents and we persuaded them to divestment policies [from fossil fuel]. We have all sorts of projects in California. I worked with 350 Bay Area (a non-profit climate action advocacy group). We worked to ban single-use plastic. Did you know that we make single-use plastic out of fracking byproducts? We didn't get the bills out this year but we will next year.

We all need to try to make a difference. But everyone needs to get out there and try to make an impact. People should join an organization or a club that advocates change. These fossil fuel industries intentionally mislead and they fight dirty. But the first step for people is to join a group, draft a strategy, and fight. We demanded that the [UC] Regents divest officially, and it worked. More like this can be done. It's going to take a lot of work from everyone.

Interviewer: That's all the time we have today. I'd like to thank you for your time. It was a pleasure!

Prof. Brown: You're welcome!

Photo Credit: KINAXIS BLOG

Professor Stephen Pratten

Interviewed by Grace Jang

Interviewer: Hi Dr. Pratten. Could you begin by telling us about your background and early experiences that influenced your interest in economics?

Prof. Pratten: Hi Grace. Okay, sure. I've been here at Kings for quite a few years now—I started here at Kings in January 1999 and I've been here ever since. And I've been teaching on various modules as well as doing research. Before that, I was based in Cambridge University, at a research centre called the Centre for Business Research, where I was doing work on structural changes in the media industries. So that's my immediate academic background. In terms of why I came into the economics in the first place—well it's always difficult to pinpoint a specific set of reasons—but I think it had to do a lot with the context of the UK in the 1980s where I was growing up. That was a period of substantial conflict in the UK, where there were very different views about how economic policies should develop, and indeed how countries should be governed. It was a period where Margaret Thatcher was in power, and there was a lot of reaction to her policy. She was very much in favour of free market, and a lot of the debate over her policy was centred around economic issues. My interest in economics stems from that period, I think.

Interviewer: I see. Were you in middle or high school during that period?

Prof. Pratten: Secondary school—so ages about 11 to 18.

Interviewer: Great, thank you. My second question is, what led you to decide that you want to pursue a PhD in economics and become a professor, as opposed to other career options in economics?

Prof. Pratten: Certainly by the time I finished my Master's degree—I did my Master's degree at Cambridge University—certainly by the end of completing that program, it was clear to me that a very important issue was, "why was the economics discipline in such a mess, why was it in such disarray, and why was it a discipline that found it difficult to develop powerful

explanations for social phenomenon." And it seemed to me there was a very important project to be undertaken exploring the methodological limitation of mainstream economics, and it seemed to me that that was a very worthwhile project to pursue, and that really led to my interest in ontology. There were a number of researchers at Cambridge working on ontology and its relevance to understanding the current state of the economics discipline. So I was very much encouraged to apply to do a PhD there, and from then, after completing PhD, it was very much a career that I wanted to go into.

Interviewer: Could you elaborate a little more on ontology and what led you to choose it as your subfield?

Prof. Pratten: Yes, of course. A very important area of research is the methodology of economics and the area of ontology. Economics is a discipline that is very splintered—there are all sorts of different perspectives within the economics discipline—but there is an enormously dominant perspective, which is defined really at a methodological level. Mainstream economics is characterized by an insistence upon mathematical modelling, and the only way to understand the damage that conventional economics does and why conventional economics is such a limited approach to consider it from an ontological perspective.

The ontological presuppositions of the mathematical modelling methods are actually very extreme—they presuppose that social world is made up of isolated atoms. And in fact, if you look at our best accounts of the social world, it's transparently the case that social world is not made up of isolated atoms, but of complexly structured beings existing in communities which are characterized by social conventions and social rules. Once you approach conventional economics from this ontological standpoint, it becomes obvious why it has made so little progress: basically, conventional economics ends up distorting the social world so that its methods can be applied, but that's no way to make progress in developing explanatory, powerful social theory. So it seems to me that an ontological perspective on the current state of economics is vital if we are to make any

kind of progress on economics. In terms of research priorities, it seems to me that that focus on ontology is quite crucial and very important.

Interviewer: Just to touch upon that a little bit more, what was the point where you started thinking that conventional economics is very problematic and there needs to be something done about it?

Prof. Pratten: Well, one of the things that struck me even as a graduate student doing my Master's degree was how different the current conventional economics' approaches were from the kind of works that I found most convincing. I read Keynes' chapter 12 of The General Theory, I read Marx and his discussion of labour process, and I was very interested in various institutional economists like Veblen and John Commons. Now all of these approaches seemed very different from the conventional economics that I was studying during my Master's course, and a great deal more insightful. These alternative perspectives seemed to be doing a much better job of generating interesting and powerful social theory. So there was this kind of mismatch between what I felt to be the most valuable kind of economics and what was most dominant in the discipline at the time, and that really stimulated a lot of interest in trying to understand that better—why was it that the mainstream approach that dominates the economics discipline is so weak when it comes to developing powerful explanations.

Interviewer: Could you briefly tell us about your empirical research into structural and regulatory change in broadcasting industries?

Prof. Pratten: When I was in the Centre for Business Research at Cambridge, I got involved in a very empirical project that was analysing structural change in media industries during the 1990s. We were interested in looking at various kinds of quasi-market reforms within the broadcasting sector especially. During the early 1990s in Britain, there was an attempt to introduce market-like mechanisms into the broadcasting sector as a way of trying to improve the functioning of the

broadcasting sector. For example, there was the introduction of quota whereby the large broadcasting institutions like BBC had to outsource 25% of their programming to small independent program producers. The hope was that that kind of quasi-market reform would stimulate a great deal of improved performance, and what we were interested in exploring was considering how effective that was—did the quasi-market reforms actually produce improved performance in the broadcasting sector, or did they generate a lot of inefficiencies and problems?

Interviewer: What inspired your research into this?

Prof. Pratten: The reason why I got involved in this project was really because I was located in that research centre, and that research centre had obtained funds to carry out that research. We spent three years studying those quasi-market reforms and structural changes in the media industries—which I found very valuable. My main research since my PhD has always been methodology of economics and in particular ontology, which is quite an abstract field of study. So it was quite a change and quite a challenge for me to look at more concrete questions, and it was interesting for me to develop skills necessary to conduct that kind of empirical research. For example, we did a lot of case studies looking at particular firms and did extensive interviewing with key participants in the industry, whether that be commissioners within broadcasters or people at key regulatory agencies in the UK that had a remit to consider broadcasting. We also looked at government documents like government reports about the state of the broadcasting sector.

Interviewer: After you completed this three-year project at Cambridge, did you continue on this kind of research into broadcasting industries?

Prof. Pratten: We visited those issues a number of times. We did follow-up studies and looked at further developments in the sectors.

Interviewer: Thank you. My next question is, is there any other

area that you're currently looking into, and if so, how are you planning on collecting your data?

Prof. Pratten: I mean, a particular interest that I've always had alongside these philosophical interests is in the history of economic thought. My most recent research really relates to the theory of money developed by an Old Institutional economist called John Commons. He developed a very interesting and sophisticated account of the nature of money back in the 1920s and 1930s. And I don't think he has received the attention that he deserved. So my most recent work has been to re-examine John Commons' work on the nature of money. As for collecting data, it'll mostly be a close reading of Commons' own texts. He published a series of very important books, including Institutional Economics and The Economics of Collective Action.

Interviewer: For my last question, may I ask you which approach you find most convincing out of all the alternative perspectives to conventional economics like Old Institutionalist and New Institutionalist thoughts?

Prof. Pratten: That's a good question. I'm not sure if I can choose one of them really. I think there's insight in many of the alternative approaches. I mean, I find a great deal of insight in John Maynard Keynes but also there's a huge amount of insights to be obtained in close readings of Marx as well as of the Old Institutionalists like Veblen and Commons. I think what's especially interesting is that while at the substantive level these heterodox economists have different theories, if you go to an ontological level—if you go to an analysis of their commitment at the level of nature of social reality—they actually have a lot in common.

Interviewer: Thank you very much for your time.

Prof. Pratten: Thank you, it's been a pleasure.

Researcher Wei Guo

Interviewed by Vanessa Thompson

Interviewer: Hello, thank you so much for joining me today. I was hoping to talk about your research here at Cal. I know you are working in the Department of Agricultural Economics. What specific research projects are you currently working on?

Guo: Hi Vanessa. I would say I have broad research interests. However, my main research interest is in agricultural economics. Currently, I am working on two projects. One of them analyzes how land certified as critical habitats affect home prices surrounding the habitat. For example, an area will be deemed a critical habitat if endangered species are found on the land. After certifying it a critical habitat, the use of the land will become limited, restricting building and any harmful activities to the ecosystem. Because of the limited use of the land, the land value may decrease. However, properties around the critical habitat increase since having a protected natural area is desirable.

Interviewer: What goes into determining if a property is more valuable when deemed a critical habitats?

Guo: The way of determining these values is through an econometric method called regression discontinuity. We look at the houses and land surrounding the border of the critical habitat. We assume that the land one kilometer from the boundary doesn't have much difference from the value of the critical habitat. We look at the land near the boundary and then compare the value.

Interviewer: Can you tell me more about how you use regression discontinuity?

Guo: Regression discontinuity is a method with the basic assumption that there is a discontinuity, and on both sides of the discontinuity, the data is fairly similar. With regression discontinuity, it should run smoothly, except for one significant jump. In our research, the big difference is the boundary of the critical habitat. So if there is no discontinuity, we shouldn't see any large change in price. With this dataset, we do see a jump that we found significant.

Regression discontinuity is also a very popular method in economic research because it is very clean and intuitive. It is very easy to see a result just by cleaning the data and doing some descriptive analysis.*

Interviewer: Regarding other work in this field, has there been much research investigating property values and critical habitat status, or is this relatively new in the field?

Guo: Well, in terms of this dataset, not really, but I am doing one other research project with this same dataset. The data is from Zillow, which is the largest real estate database. They gave us access to their dataset, and I am using it to also look at how political propensity affects the price. So in the most recent election, we expected a democratic win, but the Democratic party lost mainly because some key states swung red. Six states supported the Democratic Party in the previous election and also swung red in 2016. I want to look at the boundary of these states before and after the election and see if there are any changes of the home price.

Interviewer: Will you be looking at state or local counties in terms of their politics and home values?

Guo: I am hoping to look at both, but start at the state level. Looking at the state level will take me maybe a month to analyze. I am planning to look at the country level, but that will be a future step to take.

Interviewer: Why do you believe the politics would affect the home pricing?

Guo: There are many potential effects. One very obvious effect is that the Democratic and Republican parties have very different stances on property taxes. The Republican party strives for lower property taxes for high-income households as opposed to Democrats who push for lower taxes for low-income households. So if a state switches to vote for a different party, property owners may believe their property taxes will change. Second, people with strong political affiliations may want to

move if their state votes for a party they do not support. So that is another possible issue that could change pricing.

Interviewer: What about the recent phenomenon that wealthier urban individuals are Democratic and live in cities, whereas rural areas are now mainly Republican?

Guo: That is a good question, and this is why we look at the boundary between the two states. To address this, regression discontinuity becomes applicable once again. I assume near the edge of the property there will be a difference between wealth and other factors. This analysis will account for these differences.

Interviewer: What about other political policies that may be influential? For example, could policies dictating the amount of welfare, the number of companies in the area, and government programs have an effect? How could those affect home pricing and where people want to live? Could other factors like how good a school district is also affect where people live and the demand for those homes? With the politics influencing those other economic factors, could that influence your results?

Guo: Right, that is why we are looking at the states and counties that swung to a new party. So there shouldn't be a big difference. I am looking at the flipped states, in which there shouldn't be a big difference. But all of this is preliminary, and we will address these as we investigate further.

What we have found so far is that before the election, the gap between the home price of the Democratic state neighbor and the swing state was small. After the election, the gap widened. This gap warrants further investigation. I speculate that the expectations on taxes and other policies changed, and may have influenced prices or influenced people to move to other states.

Interviewer: This is a fascinating topic. Studying how we value properties that protect nature and how politics affects the

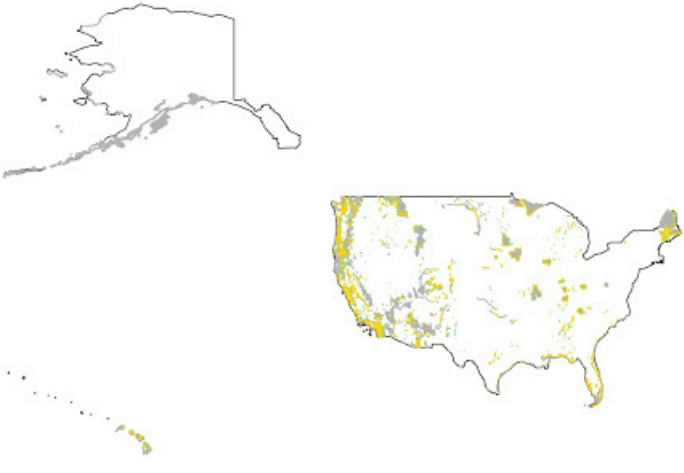
price of our homes properties are really important to investigate. Thank you so much for your time and sharing your work with us.

Guo: Of course. Thank you!

Further Discussion by the interviewer:

Natural landscapes have profound effects on communities, including making people happier, less violent, and even more generous. Wei Guo, an environmental economist and researcher at the University of California, Berkeley, is currently investigating the value of natural landscapes and their property values when protected.

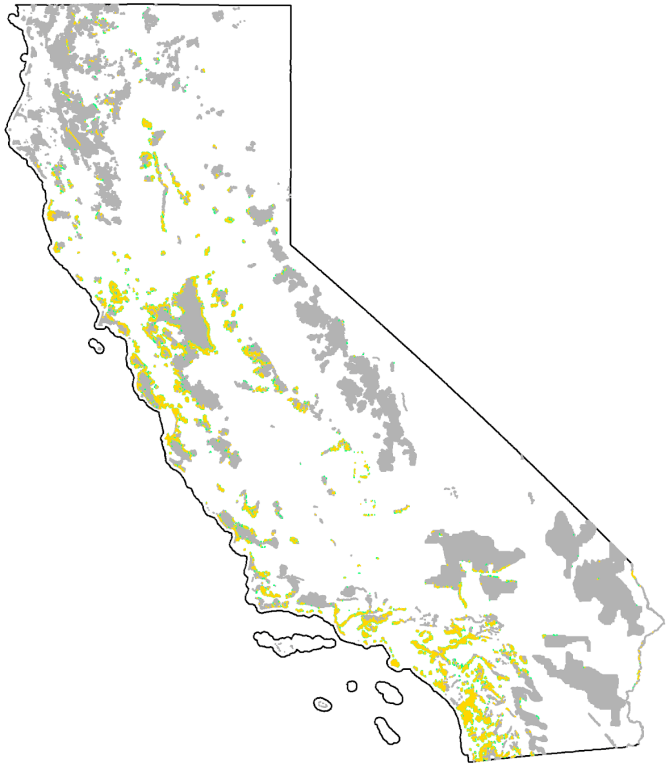
Guo's work is significant due to the large number of critical habitats. Many properties have been declared as critical habitats, which protects and provides safe areas for 704 of the 1500 endangered species in the United States. Guo's map below shows the number of properties surrounding critical habitats. The gold points show the land transactions located within 500 meters near the border of critical habitats, and the green points show the transactions within 1 kilometer. Guo's work



and the research of other environmental economists are vital to the understanding of how capital value plays a role in the protection of our natural spaces.

The figures below show the number of critical habitats in the United States and California. These show how many areas are deemed vital for ecosystem and species protection. Guo generated both using the Zillow dataset. We thank Guo Wei for generously sharing these visuals with us.

State: California. Year: 2019



PhD Candidate Mathieu Pedemonte

Interviewed by Yash Rajwanshi

Interviewer: First of all, thank you Mathieu for your time. Do you want to first introduce yourself and just talk about your background in Economics and why you started studying Economics here at UC Berkeley?

Pedemonte: Hi Yash! My name is Mathieu Pedemonte, and I am a fifth year PhD student in the Economics department. Before coming here I did a Masters and B.A. in Economics at the University of Chile.

Interviewer: What drew you to Economics specifically?

Pedemonte: I was first drawn to Economics because I liked Math, History, and Trade, and I thought Economics was a good match for that. As I worked more with it, I started liking research. Basically, the first time I had a question that was related to Economics and I saw how people had conducted research to answer that question, I was interested. I really liked this and it lead me to a PhD.

Interviewer: What diverted you to research, specifically? Why pursue a PhD?

Pedemonte: The Masters that I did was very research-oriented. To study in America from South America, you usually have to get a Masters, and the Masters that I did was very oriented to those who want to go to graduate school or PhD after. Because of that, the process was relatively easy, because my B.A. gave me an idea of my liking for research; then for the dissertation of my Masters, I worked with professors and I liked that. I slowly worked towards it and so far I am enjoying it, so it was a good decision.

Interviewer: Can you talk a little bit about your current papers that you're working on?

Pedemonte: For the last year I've been working on my job market paper—the main paper you have to develop to get a job after PhD. My field is micro or empirical macro, and what I do is work with various professors. I looked into how com-

munication can be used as a policy tool. So we know from basic economic models that expectation of what people feel will happen to the economy or "consumer sentiment" decides investment decisions, consumption decisions, and stuff like that. We started working on this idea and we realized there's not a lot of work that looks at how communication strategies of politicians making a big announcement of a policy that's going to affect the economy affects consumer behavior today. Therefore, I started answering this question and looked at some events I could use to answer this question, which lead me to a radio announcement President Roosevelt made. So I moved into Economic History, trying to understand what happened in the Great Depression and used that setting to answer this bigger question I have about how announcements affect consumers and financial decisions.

Interviewer: Referring to your previously published writings online, what are some trends you have seen that caused changes in consumer behavior?

Pedemonte: President Roosevelt used the radio for different reasons. First, he faced opposition from the media and newspapers of the time, so he wanted to communicate his ideas directly. He also used it to validate and announce the policies that he was making. In this case, the main speech that I looked at is of him announcing the Social Security Act, and other programs that were part of the second wave of the New Deal. The idea was to provide protection and stability to households. It makes sense that having that stability from the state pushes you to start investing or begin making long-term economic decisions. This was important in that time because the economy was recovering, so we didn't have that much uncertainty about the state of our economy. So I feel like this policy is interesting because it gives you some certainty of the future. Also, announcing large government expenditures have different effects today. Conventional policies for traditional banks, such as Goldman [Sachs], include their main way of stabilizing the economy, which is usually by controlling their policy rate. Today, the policy rate is relatively low and many governments have a high debt to GDP ratio, so if there's a crisis or

recession tomorrow, we won't have traditional methods to stabilize the economy. This leads to more research about other policies we can use to recover our economy. One idea is to have a future tax on goods, called the conventional fiscal policy. You announce that you're gonna increase taxes in the future, so people spend more today, providing a small boost in the economy that helps it recover from a recession. The speech I looked at talked about social security and taxes, which is why I study that specific event to consider how our learnings can be applied today based on how it worked for the government back then.

Interviewer: Some of your other research includes automation in global markets. Can you talk about that, why it interested you, and the methodology?

Pedemonte: My other focus was international trade, and my advisor and I started thinking about the idea of automation and how it can affect trade. The problem with automation is that many of the papers that look at automation think of it as something that is exogenous. They believe there is some sector in the economy that automates because of an exogenous reason, and for that reason they are affected more by automation. We believe this might actually have more of an endogenous reason. For example, if you are a manufacturer in the US and want to outsource a part of your manufacturing process, the original way was to outsource to an underdeveloped country. But today, you can create machines to increase efficiency. That way, trade has roles in automation trends. We tried to make a model and test some implications of the model with data, trying to incorporate the trade mechanism that can drive automation. We are still working on this research, even though it has paused for a bit.

Interviewer: You mentioned you're wrapping up your PhD and then want to work. Do you want to go into research or into industry?

Pedemonte: In my case, I want to find something more related to research. I like the research process and I have some re-

search ideas and working papers that I want to develop. For that reason, I want to give a shot to my current projects to see if they are good enough to get published and create a career off of. The research market is unstable and even professors will tell you that you don't know where you will end up for research. Right now I am looking at universities and those type of jobs, but also there are some central banks and international organizations that have research roles, so I am also applying to those jobs. However, my main idea in the next few years is continuing my research.

Interviewer: After all your years studying Economics, what advice do you have for a student getting into Economics right now at UC Berkeley?

Pedemonte: I always say that Economics is fun, because I didn't have Economics in high school and we still had big questions about inequality and the function of the government or political decisions. Economics gives you a lot of tools to answer and have an opinion about these things, and for me, I started understanding what was going on and I really enjoyed that. My first advice is enjoy Economics, don't think about problem sets or specific lectures, but think about the implications of them and how you can develop an opinion around what you're learning. Also, at Berkeley, you have access to really good professors that are talking about new things that are super relevant. You have professors that are talking about minimum wage, monetary policy, and all the other key current events in Economics. Look at professors that are interesting and take their classes and ask questions. Being a curious student, even if you don't want to do research, is useful here at UC Berkeley because of the amazing faculty group. Even if you don't want to do research, you should use your toolset to make a change in your community. Unfortunately, sometimes economists are valued too highly in policy decisions. However, if you understand Economics because of your education, you can use this power to impact your community. For example, AOC [Alexandria Ocasio-Cortez], who studied Economics, is bringing up basic economic concepts and that leads to people taking her more seriously. Making the link between what you're learn-

ing and the world around you is important given how relevant Economics is to all the decisions in the world.

Interviewer: Awesome, thank you so much for your time Mathieu.

PhD Candidate Zachary Bleemer

Interviewed by Konnor von Emster

Interviewer: Zach, thank you so much for meeting with me today. To start off, would you mind giving our readers a quick synopsis of your education and research thus far?

Bleemer: Sure—I did my undergraduate at Amherst College. I had three majors: in economics, but also math and philosophy. I then went and worked at the New York Fed [Federal Reserve] for 2 years. I was primarily working on a project that had to do with student debt. The massive accumulation of student debt—the impact of student debt—for post graduate outcomes for college students, in particular. We were interested in students who were moving home with their parents after they graduated. We also thought that this also had ramifications for the housing market because so many [new graduates] were moving back into their parents' homes instead of purchasing a new place. So I spent two years working on questions like that for the New York Fed, then came here as a PhD student.

My current research agenda examines the long-run ramifications of a long variety of university policies, primarily focusing on university admissions, but also looking at major choices and interactions between students and professors. This work allows me to get a sense of how these policies shape the long-run outcomes of young people who are making extremely important choices with very limited information, and it turns out that which university you go to or which major you choose is extremely impactful of the kind of life you live after you leave university.

Interviewer: What inspired you to specifically dive into that field? The one of regarding education equality in that of student loans, affirmative action, ELC, etc....?

Bleemer: I came out of undergrad very interested in how people decided what to do with their lives. I saw my peers making big choices, and while some choices seemed good, and others did not. I just became interested in analyzing what mattered, and what didn't, in these decisions made by young people and what the repercussions of those choices were. Naturally, I became interested in policies that forced people to choose

one way or another, specifically by asking: "what would have happened if people made a different choice?" For instance, race-based affirmative action is a policy that provides access to selective universities to disadvantaged groups of students. I think this is interesting for two reasons. One reason is on the equity side; these are students who otherwise would not have had access to tremendous educational resources, so as universities phase out race based affirmative action policies, they contribute to heightening social immobility in the United States. Secondly, these policies are interesting because they let you ask the question: "what would happen if people didn't have access to these selective universities?" I have a lot of friends who chose against going to such selective universities. I come from rural Pennsylvania, and most of my friends went to local universities, so affirmative action was a way of studying two groups of people: those who get into selective research universities, and others who don't. After this division, you can follow them along and see what happens to people who make the selective university choice. Overall, [I think] Affirmative action is very positively impactful to targeted students, as university enrollment is very economically advantageous for young people.

Interviewer: You looked a lot at the mismatch hypothesis in that study; can you explain a bit more about that?

Bleemer: Thomas Sowell is a prominent, conservative economist, who in the 1970s, wrote a book about the ways in which affirmative action could be harmful to the black and Hispanic students that it targets. He makes a number of arguments, and the central argument of which is the mismatch hypothesis: the notion that taking a student with a low measured preparedness—someone of which has low standardized test scores for instance—and admitting them and ultimately encouraging them to enroll at a selective research university where they will really be different than their peers in the sense that they will have lower test scores, often have lower income. Sowell was concerned that this would be, in the long run, damaging to these students because they would have a difficult time competing in their courses and they would get pushed out of

lucrative and difficult fields. They would also be less likely to graduate because they are having a hard time getting by in college, and he thought these could lead to long run labor market declines for targeted students. This has recently been picked up by Richard Sander, who is a law professor and economist at UCLA, who has been a strong proponent of ending affirmative action policies at selective universities. And my research suggests that this hypothesis is false. In other words, the notion that low-testing black or Hispanic students who enroll in selective research institutions as a result of access-oriented admissions policies like affirmative action might have been harmed is false. It turns out that they're likelihood of graduating increases under affirmative action. Although there's no measurable impact on their likelihood of earning lucrative, challenging majors in STEM fields, for instance, their long run labor market outcomes do improve.

Interviewer: Do you think it's almost harmful to take up these affirmative action policies or is there a better solution, such as "Eligibility in the Local Context?" Or can you speak to the ideal solution?

Bleemer: Let me first try to quantify the ramification of ending affirmative action. The University of California has not used race based affirmative action in their undergraduate admissions since 1998. If you compare long-run labor market outcomes for black and Hispanic applicants to the university in 1997 (when they still had affirmative action), verses 1998, what you see is about 700 students per year who no longer attended the University of California at all. Based on the absence of race-based affirmative action, there are far more students who got pushed down the selectivity rungs within the University of California. They were previously able to go to Berkeley, but now they go to San Diego, etc. What that means is 15 years later, when these students are in their early 30s, California has about 20,000 young black and Hispanic workers who earn more than $100k/year; these are a wealthy group of underrepresented minorities, composed of high-earning young people. However, because we ended affirmative action, that group declined in size by about 3%. Call it 600-700 people who would've been

high earning black or Hispanic workers in California, who now weren't because we ended affirmative action. The ramifications are modest but real.

Interviewer: And it's certainly not the opposite trend, as suggested by the mismatch hypothesis.

Bleemer: Yes, that's certainly important. To give you a sense of scale, ending affirmative action was harmful to these targeted groups. You brought up Eligibility in the Local Context (ELC), which was the University of California's attempt at a race-neutral response to ending its affirmative action program. The idea is that the university would guarantee admission to 4% of students from every high school in the state of California. Because of dramatic heterogeneity across California high schools, you can imagine that there's good high schools and bad high schools largely determined by local income. At good high schools, the top 4% of kids could already get into a UC school and the program does nothing; but at a bad high school, the program is extremely impactful. So at these low high schools, this was targeting students who had already enrolled at a local community college and sending them a letter saying: you've been guaranteed admission to the University of California. So this was a radical change for those impacted students. And again, as with the case of affirmative action, what happens to those students? They become more likely to earn college degrees, more likely to enroll in graduate school, and their early career earnings are substantially higher. In fact the benefits to students scale in regards to the negative quality of their counterfactual. The worse the school the student would have otherwise gone to, the greater the benefit of getting pulled into these selective public research universities in California. That's exactly the opposite of the mismatch hypothesis, which states the poorer the fit of the student, the harder time they're going to have at a better university. In fact, the benefits positively scale for the low quality students. Your question of whether one policy or the other is optimal for the university is fundamentally a question of tradeoffs. There are a lot of types of disadvantaged young people and they all benefit from selective research universities, or at least many of these

groups. It's a policy choice of who to allocate selective university enrollment to. Both of those policies were largely impactful, ELC a little bit smaller than Affirmative Action but still quite large at 100s of students a year. It's not the role of a researcher to say who should be given educational resources; all that I can say is that both of those groups were extremely positively impacted.

Interviewer: Do you think this research could inform other states and selective universities' admission policies?

Bleemer: Yes, absolutely. Affirmative action policies and other race-based affirmative action policies are waning as a result of political movements and judicial actions. Public and private universities are now looking for alternative ways of targeting disadvantaged students and providing educational resources to them. And very much motivating this work I've done in Eligibility in the Local Context in California has been to try to lay out the advantages of this particular alternative program, guaranteeing admission to students in each high school in the state. One state was already using this admissions policy before California, which was Texas. Florida and Georgia have implemented similar policies since. A lot of other states are looking at whether to implement this kind of policy and yes, I expect that this work will be helpful to them to gather a sense of whether this policy would help the students they're targeting and the universities they're affecting.

Interviewer: Are there any other similar policies that you are looking into?

Bleemer: The next policy that I will focus on and have done a little bit of work on is "Holistic Review." So this is a very commonly used admissions policy at public and private universities; the idea of this policy is to not use fixed weights across applicants' applications in determining who should be admitted to the university, but instead looking at the whole student. Since there are many components of the [college] application, the policy will make a choice based on all aspects in combination to determine applications. Holistic review looks in the

data a lot like an affirmative action program, in that switching to a holistic review policy substantially increases black and Hispanic admissions, as well as students with low household income and other socioeconomic disadvantages.

Interviewer: Just to clarify, this is a switch from a non-holistic review process that purely looks at test scores and GPA?

Bleemer: That's right. Prior to holistic review, the universities I've studied still ask for the full application, but have a fixed set of weights across the components. 15% is test scores, 10% is high school grades, etc.... And there's relative inflexibility of an essay overriding the information gained from test scores. In holistic review there is more flexibility; you can learn about a student's disadvantage from their essay, or access to educational opportunity, and reweight a student's test scores with that new information. The last thing I will briefly say is that the University of California is considering eliminating the SAT from its application. That would be a big choice that will likely be made next year. And that would act as an access-oriented admissions policy. If you're not able to target admissions towards high testing students, then the number of disadvantaged applicants would increase. So that's another policy that's of interest to some policy makers.

Interviewer: Well, that seems to be all the time we have. Thanks so much for talking with me!

Bleemer: Of course Konnor. It has been a pleasure.

Interviewer: How did you convince the universities to give up so much data?

Bleemer: This is a really good question for an undergraduate audience. The hardest part about writing a research paper as an undergraduate is finding data. I work with a lot of students writing theses or taking Econ 191 and that is the central challenge they face. People are able to develop technical skills and are completely ready by senior year to solve econometric problems but they have no data to work with. The key thing

you need to do in order to collect data is to find people who have data and do something for them. And it's often difficult to figure out what you can do for an administrator, in my case, a university administrator. What I did for them was collect this historical information of the university that the administrators did not have, but interested them. They could tell interesting stories about the university which was good PR for the university. Everyone wants good PR, no matter the organization that holds data. You can make this trade. Offer them information they do not have and in return universities have been willing to provide contemporaneous student records. It's always a tradeoff: I do work for them, they're willing to provide things to me.

Interviewer: Thank you so much for speaking with me. It has been a pleasure.

The Impact of Knowledge Economy Factors on Total Factor Productivity: Evidence from the Asian Leaders

Laure Fleury
Maastricht University

Abstract

This study provides an insightful examination of the contribution of the four pillars of the knowledge economy to Total Factor Productivity (TFP) improvements in a panel of six leading Asian knowledge-based economies — Hong Kong, Japan, Singapore, Republic of Korea, Malaysia and Thailand — over the period 1996-2017. Based on a panel ARDL-PMG model, the results appear relatively fragile. Nevertheless, establishing upon the most recurrent relationships, it appears that domestic innovation, education levels and the access to ICT are important drivers of *TFP* enhancements.

Keywords

Total Factor Productivity, Knowledge-based Economy, East Asia

Table of Contents

Acronyms and Abbreviations

ADB	-	Asian Development Bank
ARDL	-	Autoregressive Distributed Lag
ASEAN	-	Association of Southeast Asian Nations
COFOG	-	Classification of the Functions of Government
EBA	-	Extreme Bounds Analysis
ECT	-	Error Correction Term
FDI	-	Foreign Direct Investments
GERD	-	Gross Domestic Expenditure on R&D
HPAEs	-	High Performing Asian Economies
ICT	-	Information and Communication Technology
IT	-	Information Technology
KAM	-	Knowledge Assessment Methodology
KE	-	Knowledge Economy
KEI	-	Knowledge Economy Index
LICs	-	Low-Income Countries
NICs	-	Newly Industrialized Countries
OECD	-	Organisation for Economic Co-operation and Development
PMG	-	Pooled Mean Group
R&D	-	Research and Development
TFP	-	Total Factor Productivity
UNCTAD	-	United Nations Conference on Trade and Development

1 Introduction

The East Asian economic success, or what the World Bank (1993) defined as the "miracle," has largely been investigated in the literature. There have been many attempts to offer explanations on the remarkable growth of Japan in the 1950s and 1960s, the Four Tigers (Hong Kong, Taiwan, Singapore, and the Republic of Korea) subsequently, and more recently three of the ASEAN-4 (Thailand, Malaysia, Indonesia). Based on the growth accounting findings of Young (1994) and Kim and Lau (1994), Krugman (1994) discovered that hyper-growth in East Asia had been driven by factors input accumulation—the growth in quantity of factors of production like capital, labor, land etc. These findings led to a certain scepticism concerning the futures of the aforementioned countries' economies, which could fail to sustain their economic prosperity in the future. In fact, the accumulation of factors of production is subject to diminishing returns and, hence, is by itself insufficient to achieve long-run growth. An upsurge in productivity is required to yield higher output with the same amount of resources to generate increasing returns.

Recognizing the necessity of higher productivity levels, significant research has been conducted on exploring the effects of Total Factor Productivity (*TFP*)-led economic growth and its determinants. *TFP*, often referred to as technology or efficiency, is the Solow residual of output which captures everything that is not explained by the traditional input factors (Solow 1956). The Solow model was the first model in economic growth literature which recognized the importance of *TFP* as a sustainable driver of economic growth and social welfare improvement. However, although *TFP* represents a central component of neoclassical theories, it is left unexplained and assumed exogenous. Endogenous growth theories have sought to understand what is in this black box that neoclassical theories have classified as exogenous. Romer (1986) and Lucas (1988) emphasized the importance of knowledge accumulation, both through the production of ideas as a by-product of work experience, as

well as from foreign knowledge spillovers, as the main source of sustainable growth.

A decade later, the 1998/1999 World Development Report accentuated the major role of knowledge in advancing economic progress and sustaining social well-being (World Bank, 1999). The four pillars of successful knowledge-based economies were identified. These economies are (1) governed by a supportive economic and institutional regime which encourages the establishment of (2) an effective innovation system, (3) the access to information and communication infrastructure, and promotes (4) the accumulation of knowledge and skills (World Bank 1999). In the aftermath of this publication, the Asian Development Bank (ADB) held their 39th Annual Meeting, "Innovative Asia," to encourage the transition of Asian countries towards knowledge-based economies (Asian Development Bank, 2007). However, proof remains fragmented as to whether the accumulation of knowledge has indeed fostered *TFP* improvements in the Asian knowledge-based economies. Has that potential been truly realised? This study investigates the relationship between the four Knowledge Economy (KE) pillars and *TFP* improvements in the leading Asian knowledge-based economies over the period 1996-2017.

This paper is organized as follows. Section 2 presents an overview of the previous empirical research. In Section 3, the research design is described: first, the motivation and relevance of the study are presented; second, the data compiled and indicators selected are described; third, the econometric specifications of the model are presented. In Section 4 and 5, the empirical findings are interpreted, and the underlying policy implications are suggested. Section 6 presents the limitations of the study and suggestions for future research. Finally, Section 7 concludes this study.

2 Previous Research

With the advent of new growth theories, there have been a considerable number of studies that explore the determinants of *TFP* and *TFP* growth in East Asia. In this section, I review previous empirical studies which have considered some of the four KE pillars and their relationship with *TFP* in East Asia. Since abundant research exists, I present the broader lessons of the literature narrowed down to the most recent findings. It is clear that far more can be learned than I can convey here.

One of the pillars of the KE concerns the innovation system. Coe and Helpman (1995) constructed a theoretical model which presents innovation-driven growth. Based on this model, several studies have demonstrated varying findings concerning the significance of the innovation system in East Asia. For instance, Madden et al. (2001) compares the effect of domestic and foreign Research & Development (R&D) capital stocks on *TFP* growth in 16 Organisation for Economic Co-operation and Development (OECD) countries and 6 Asian economies between 1980 and 1995. They observe that domestic R&D expenditures have had a strong, positive impact on *TFP* in Newly Industrialized Countries (NICs) (Chinese Taipei, Korea, Singapore) and Asian Low-Income Countries (LICs) (India, Indonesia, Thailand). The elasticity of *TFP* with respect to domestic R&D in this group of countries is more than three times higher than for the OECD countries. Furthermore, their results suggest that the impact of international R&D activities is mixed. On the one hand, foreign R&D spillovers are significant and strongly positive in Japan, Taipei, Indonesia, and Thailand. On the other hand, the spillover effects are not significant for Singapore, Korea, and India. These findings are disproved by Okabe (2002) who observes that, in the same period, R&D expenditures in OECD countries have promoted higher *TFP* levels in Singapore, Hong Kong, Indonesia, Korea, Malaysia, the Philippines, and Thailand. Moreover, the author adds that international R&D has been highly significant through the channel of trade, and most

specifically via importing manufactured goods. In fact, this supports the review of Isaksson (2007) which indicates that Foreign Direct Investments (FDI) and trade are the two major channels for knowledge transmission. Evenson and Singh (1997) show similar evidence for the strong, significant effect of international knowledge spillovers in encouraging *TFP* growth in 11 Southeast Asian countries over the period 1970-1993. However, Ang and Madsen (2011) criticize the R&D-led growth theories in the case of High Performing Asian Economies (HPAEs). They come to the conclusion that there has not been any long-run relationship between domestic or foreign R&D and *TFP* growth between 1953 and 2006. Instead, the authors prove the existence of a relationship between R&D and product variety, effectively supporting the Schumpeterian view.

The accumulation of skills and education is another fundamental component of the KE. According to endogenous growth theories, human capital is the crucial determinant of the capacity of technological innovation (Romer, 1990), but also governs knowledge absorption and technology adoption (Nelson and Phelps, 1966). Therefore, in order to assess the significance of R&D expenditures and R&D spillovers, multiple studies have investigated the importance of education levels and skills. Mahmood and Afza (2008) analyze different knowledge components as determinants of *TFP* growth in East Asia between 1980 and 2000. The authors find that, among various indicators of education, domestic R&D activities, and foreign R&D spillovers, only secondary education has a significant impact on *TFP* growth. These results are in line with the findings of Zachariadis (2004) on the importance of secondary enrolment in fostering *TFP* growth in OECD countries. In a comparative study of Asia and the OECD, Park and Park (2010) discover that human capital, as proxied by 5-year averages of educational attainment levels, has been a critical driver of *TFP* growth, especially in Asian economies. The authors further indicate that investment in human capital leads to higher levels of openness and more efficient governance. The positive impact of human

capital was recently contradicted by Yuhong et al. (2017) who perform a more sophisticated analysis by allowing for a nonlinear relationship between human capital and *TFP* growth. The authors discover that in a panel of 8 ASEAN countries from 1990-2014, secondary education attainment is, by itself, negatively related to *TFP* growth. Only in conjunction with the catch-up term – the technology gap – does human capital generate a positive effect on *TFP* growth. These results illustrate the importance of human capital in effectively interacting with technology spillovers from the frontier.

The ability to communicate over large distances using technological means is an integral part of knowledge-based economies. However, previous studies have often failed to represent the correlation between Information and Communication Technology (ICT) investment and productivity improvements. According to the productivity paradox of Information Technology (IT), there are two substantial impediments to research attempting to illustrate this relationship: the incorrect measurement of outputs and inputs and the presence of lags to learning and adjustments (Brynjolfsson, 1993). This theory is often associated with the earlier quote of Solow (1987:36): "Computers are visible everywhere except in productivity statistics." Despite the progress in measurement methods and the increasing availability of data on intangible assets, studies on the impact of ICT on aggregate *TFP* levels in Asia remain scarce. Kraemer and Dedrick (1994) challenge the paradox in a study of 12 Asia-Pacific countries between 1984 and 1990. Their findings show a strong correlation between investment in computer hardware and software and productivity advancement. Moreover, they confirm that education is highly correlated with ICT, as modelled by Nelson and Phelps (1966). Similarly, Ahmed (2017) identifies the role of human capital and ICT in eight countries from the ASEAN5 and East Asia. The results indicate that productivity growth is positively influenced by ICT and human capital in the period 1965-2006.

Knowledge-based economies require an economic and institutional regime that establishes adequate incentives. The role of the government in fostering productivity growth has received increasing attention in recent years. In their findings, Park and Park (2010) identify the roles of integration and government effectiveness, although of less significance than that of human capital, as drivers of *TFP* improvements in the OECD and Asia between 1992 and 2007. These results confirm the conclusion of Rodrik (1997) concerning the importance of institutional quality in fostering *TFP* growth in East Asia. According to Rodrik, the institutional context in which government interventions are implemented is highly important. Nevertheless, the question of whether governments should intervene or not has been the subject of much debate in the literature. While some economists support the "governed market" view, or the idea that governments should intervene in the markets, others support the "free market" theory of laissez-faire governments. According to Nelson and Pack (1999), government intervention has represented a critical component of *TFP* enhancements in Asia between 1960 and 1996. Governing entities have been able to establish efficient policies and public investment efforts to foster major structural changes and allow resource reallocations. Furthermore, Thomas and Wang (1996) analyze panel data for 58 developing countries and 10 East Asian economies between 1960 and 1990. They measure two indices of government intervention. On the one hand, the index of government integration and stability is significantly correlated with *TFP* growth. On the other hand, government expenditures are a non-linear driver of *TFP* growth. Nevertheless, the authors recognize that investigations into the influence of government intervention on *TFP* is hazardous as it depends on the nature and the quality of intervention. In fact, Knowles and Garces-Ozanne (2003) argue that government spending is an inadequate proxy of government intervention as it fails to capture the non-financial public influence. In East Asia particularly, government ownership of key businesses, banks, and real estate, or the specific laws and regulations established, are

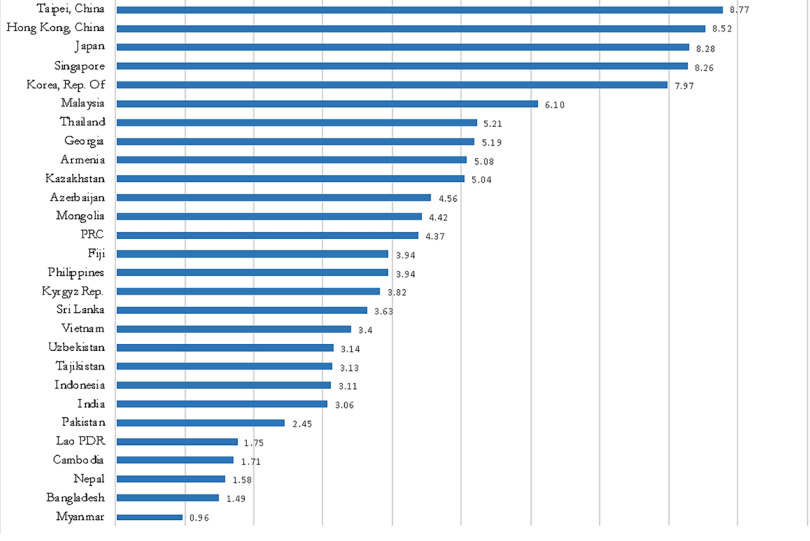
other serious ways through which governments influence *TFP*. They examine different proxies of government intervention. They find that red tape has had a positive effect on productivity, while government ownership has had a negative effect. Furthermore, there appears to be weak evidence for price controls influencing *TFP*.

3 Research Design

3.1 *Motivation*

The conclusion that can be drawn from the above section is that the existing empirical literature is highly fragmented. In fact, although previous studies have analysed different determinants of *TFP* and *TFP* growth in Asia, with amongst them knowledge variables, no study has gathered the four pillars of the KE yet. This study is therefore of academic relevance as it fills this important gap in the existent literature on *TFP*. Moreover, this study focusses on the most recent figures and analyses the most recent data so far in this regard.

Figure 1. Knowledge Economy Index Scores (2012)



Source: Asian Development Bank (2014)

Most importantly, this study is of social relevance as it aims at guiding Asian policymakers in the establishment of sustainable development measures which create social welfare. The Knowledge Economy Index (KEI) scores depicted in Figure 1 reveal that most Asian economies perform below the OECD average in establishing knowledge-based economies for economic development. Only four countries – Taipei, Hong Kong, Japan and Singapore – lie above the OECD average index of 8.25. According to the ADB, the economies on their transition towards knowledge-based economies should learn from the Asian leading knowledge-based economies which have already achieved a remarkable transition (Asian Development Bank, 2014). Because it would be financially unsustainable for those developing economies to invest into the four knowledge pillars simultaneously, governments are advised to learn from the experience of the leaders in order to identify the most efficient investment portfolio that will lead them to long-run enhancements of *TFP* levels.

In this objective, this study analyses the effect of knowledge components on *TFP* in a panel of the six leading Asian knowledge-based economies according to the KEI scores: Japan, three of the Four Tigers (Hong Kong, Singapore, and Republic of Korea)[1] and two NIEs of Southeast Asia (Malaysia and Thailand). An attempt is made to test for the hypothesis that *TFP* levels depend upon the four pillars of the KE, as mentioned by the World Bank (1999). For this purpose, the general form of the model is:

$$(1)\ TFP_{i,t} = f(TFP_{i,t-1}, \Delta INNOV_{i,t}, \Delta EDUC_{i,t}, \Delta ICT_{i,t}, \Delta INSTIT_{i,t})$$

where *TFP* in country i is assumed to be a function of past *TFP* levels and four vectors of variables representing the four pillars of the KE: (1) the innovation system, (2) education and skills, (3) the information and communication

1 - Taipei, the fourth Tiger and the Asian economy ranked first with the highest KEI, was excluded from the sample because of the lack of country-specific data distinguished from the series of Mainland China.

infrastructure, and (4) the economic and institutional regime.

3.2 *Data and Variables Description*

This study uses macro-level annual data for a panel of six leading Asian knowledge-based economies cited above. The use of panel data allows to reduce the issue of multicollinearity among explanatory variables and enables to control for the effects of missing or unobserved variables. Moreover, the panel data helps to overcome the issue of limited data availability and increases the number of observations and degrees of freedom. Regarding the time period, the main focus of this study is to analyse the period of transition towards knowledge-based economies, i.e. the most recent decades. Accordingly, data spanning 22 years from 1996 through 2017 is analysed. Next, I elaborate on each variable used in the model.

Total Factor Productivity

In this study, I measure *TFP* levels for each country. The measurement has been founded on its theoretical foundation (Solow, 1957). The framework is based on the aggregate production function. Building on the existent literature, I decide to specify the production function in its most common style, in the Cobb Douglass form[2]:

$$Q_{it} = A_{it} K_{it}^{\alpha} L_{it}^{\beta} \quad (2)$$

where Qit represents aggregate output in country i at time t, and is a function of the stock of physical capital, Kit, labour inputs, L_{it}, and *TFP*, A_{it}. *TFP* dictates the shifts in the isoquants of the production function. Equation 2 indicates that an increase in *TFP* has the effect of increasing the productivity of both input factors in a Hicks-neutral manner (Hicks, 1966). I derive the Solow residual directly

2 - Another common approach has been to specify the general production function in its translog form (Caves et al., 1982).

from the above equation:

$$(3) \qquad TFP_{it} = A_{it} = \frac{Q_{it}}{K_{it}^{\alpha} L_{it}^{\beta}}$$

Time series from the Penn World Table 9.1 are used to compute Equation 3 (Feenstra, Inklaar et Timmer 2015). The series used are the real GDP, the capital stock, the labour force[3] and the share of labour payments. Only under two crucial assumptions can I compute and with the values of labour payments data available. First, I assume that capital and labour markets are perfectly competitive and, hence, that the marginal product of each factor equals its respective price. Thereafter, I build upon the assumption that production is characterized by constant returns to scale, i.e. $\alpha+\beta=1$, to allow for the derivation of capital payments from the available labour payments series. The Solow residuals, "*TFP*", that I compute for each individual country are displayed in Table A1.

Innovation System

The innovation systems – within firms, research centres, universities, think tanks, and others – in knowledge-based economies must conduce the creation and assimilation of knowledge and new technologies. The quality of innovation systems is hardly measurable. In the existent literature, it is often proxied by the stock of R&D investments measured with Gross Domestic Expenditure on R&D (GERD) or the number of researchers. Another commonly used indicator of innovation is the stock of patent grants. Nevertheless, whereas R&D figures represent input in the production of knowledge, patent counts provide a measure of the output. Presumably, patents represent the ideas themselves and, hence, the underlying knowledge stock (Hall, Griliches and Hausman, 1986). Consequently, in this study the variable representing the internal innovation, "Innov", is based on resident patent data retrieved from the World Intellectual Property Organization (2018). Limited by the availability of

3 - I compute the labour force by multiplying the number of persons engaged per thousand individuals with the average annual hours worked.

data on resident patent grants, resident patent applications series are used instead.

Building on Griliches (1980), knowledge stocks are constructed from cumulated patent applications by applying the perpetual inventory method:

$$(4) \qquad P_{i,t} = \Delta P_{i,t} + (1 - \delta) P_{i,t-1}$$

where $P_{i,t}$ is the stock of resident patent applications in country i at time t, which depends on the change of resident patent applications at time t, $P_{i,t}$, and the depreciated patent stock of the previous period $1\text{-}\delta P_{i,t-1}$. Consistent with the literature, the patent stock preceding the initial year is generated with the formula $P_{i,0} = \frac{\Delta P_{i,0}}{g_i + \delta}$, where $\Delta P_{i,o}$ is the number of resident patent applications in the first series available, and g_i is the average growth rate of patent applications in country i. The attempt to estimate the rate of obsolescence of knowledge represents a major research theme by itself and is beyond the scope of this paper. The latest study done by Park, Shin and Park (2006) on the depreciation rate of patents estimated an average rate of 13.3% for the period 1985-1999 in the United States. Therefore, I assume a constant patent stock depreciation rate of 13.3% for all sampled countries.

Knowledge is not only created internally but may also be imported from advanced foreign countries. Two major channels exist through which knowledge spillovers can arise (Isaksson, 2007). On the one hand, a common channel of knowledge is FDI. Interactions between countries represent an important transmission of knowledge. Moreover, FDI increases the competitiveness of domestic firms for which innovation becomes vital. The variable "FDI" is measured by the inward stock of FDI retrieved from the United Nations Conference on Trade and Development (UNCTAD, 2019). On the other hand, knowledge and ideas may flow across national borders via international trade. Technological know-how may be embodied in certain goods or services imported from advanced foreign countries. Knowledge spillovers through imports are represented by the variable

"Import," measured by total imports series provided by the World Bank Database (World Bank, 2019).

Education and Skills

The level of education and skills drives the speed of technological innovation (Romer, 1990) and the degree of absorptive capacity (Nelson and Phelps, 1966). Indeed, the significance of technology transfers through both FDI and trade channels highly depends on the degree to which one can absorb new knowledge and adopt new technologies. Countries should encourage the accumulation of skills and emphasize an education of quality which enables populations to create, share, and use knowledge efficiently. The measurement of human capital has attracted considerable attention in the literature. School enrollment or education attainment have often been used as proxies. However, these metrics do not represent the quality of educational systems which varies across countries. Moreover, those proxies do not represent the outcomes of education from sources other than school such as families or trainings. Hanushek and Kimko (2000) constructed a metric that aims at representing cognitive skills, based on average test scores on the PISA test. Constrained by the availability of time series data for this metric, the variable indicator of human capital in this study, "Educ," is computed from the index considering average years of schooling (Barro and Lee, 2013; Cohen and Soto, 2007; Cohen and Leker, 2014), and the rate of return to education (Psacharopoulos, 1994) retrieved from Penn World Table 9.1. This human capital index not only represents school education, but also abilities or skills accumulated from experience at work or from external education or trainings.

Information and Communication Infrastructure

Efficient information and communication infrastructures

4 - 1980 for Republic of Korea and Thailand, 1983 for Hong Kong and Japan, 1985 for Malaysia and 1995 for Singapore.

integrate personal computers, mobile phones, and the Internet efficiently to facilitate the sharing and accessibility of knowledge. Countries must use ICT to embrace the characteristics of the knowledge economy such as openness, efficiency, and interaction, in order to allow the exchange of knowledge across borders. This third pillar is represented by the variable "ICT" measured by the population using the Internet per thousand individuals retrieved from the World Bank Database (World Bank, 2019).

Economic and Institutional Regime

Considering North and South Korea or mainland China and Hong Kong, it is clear that divergence in economic outcomes highly depends on national differences in the quality of economic policies and institutions (Olson, 1996). In a knowledge-based economy, the economic and institutional regime must introduce appropriate policies that conceive dynamic and flourishing economic environments. Institutions of quality incentivize the efficient use of existing knowledge and encourage individuals to engage in innovation and the creation of new ideas. According to Hall and Jones (1999), it is institutions and policies – the social infrastructure - that influence the nature and quality of investments made into capital, skills, and technology, and hence, drives the differences in productivity levels across countries. In this study, I follow their specification of the social infrastructure for the measurement of the fourth pillar. Indeed, I use similar series from the data assembled by the PRS Group in the International Country Risk Guide providing ratings according to 22 variables (PRS Group, 2019). Following Knack and Keefer (1995), I take into consideration the following components: (i) law and order, (ii) bureaucracy quality, (iii) lack of corruption, and (iv) investment profile, which considers the risk of expropriation and government repudiation of contracts. Additionally, I consider two extra components: (v) government stability and (vi) democratic accountability. For each index, a maximum score of 6 signifies very low political risk and, hence, high quality of institutions. In total, six of the

Political Risk Components, for which equal weight is given, are used to construct an index of social infrastructure, namely "Institu."

In Table A2, a summary of the variables along with their descriptive statistics and data sources is presented. All indicators in million dollars have been deflated to 2011 real constant values using the GDP deflator to account for the price effect. Building on the 1998/1999 World Bank Report, I hypothesize that each of the four pillars of the KE should have a positive impact on *TFP* levels (World Bank, 1999).

3.3 *Econometric Specification of the Model*

In order to investigate the relationship between the aforementioned variables and *TFP* in a sample of 6 leading Asian knowledge-based economies, the study adopts a panel Autoregressive Distributed Lag (ARDL) model based on the techniques introduced by Pesaran et al. (1999). This econometric framework is adequate for several reasons. First, the panel ARDL model is highly efficient for small sample sizes as in the case at hand (N=6). Second, this model corrects for endogeneity by including a certain predetermined number of lags—notably the lagged values of the dependent and independent variables. Third, and most importantly, the ARDL approach allows the examination of long-run and short-run relationships, as well as the speed of adjustment from the short-run disequilibrium to the long-run equilibrium. In other words, the ARDL model provides more sensitive and precise estimations, which may adequately guide policymakers. Under the assumption that long-run coefficients are homogenous across the selected countries, I decided to analyze the Pooled Mean Group (PMG) estimator of the ARDL model (Pesaran et al., 1999). This estimator, which involves both pooling and averaging, allows the intercepts, the short-run coefficients, and the error variances to vary across countries.

The re-parameterised panel ARDL (p, q, q, ..., q) error correction model is specified as follows:

$$(5) \qquad \Delta \ln ln\, TFP_{it} = \varphi_i \left(\ln ln\, TFP_{i,t-1} - \beta_i' X_{it} \right) + \sum_{j=1}^{p-1} \alpha_{ij}^{*} \Delta TFP_{i,t-j}$$

$$+ \sum_{j=0}^{q-1} \delta_{ij}^{*'} \Delta X_{i,t-j} + \mu_i + \varepsilon_{it}$$

In Equation 5, Xit is the vector of explanatory variables detailed in the previous section in their logarithmic form: *ln (Innov) , ln (FDI) , ln (Import) , ln (Educ) , ln (ICT) , and ln (Institu)* . The first parenthesis represents the Error Correction Term (ECT), preceded by the speed of adjustment coefficient i. The parameters i represent the long-run effects of the explanatory variables. The remaining parameters ij* and ij*' illustrate the short-run relationships. The constant i is the country-specific fixed effect and p,q are the optimal lag orders.

Before estimating the model, a series of specifications and diagnostics must be performed. First, in order to avoid multicollinearity, it is encouraged to practice a correlation analysis of the selected variables. The results are illustrated in Table A3. I decided to not include ln (Import) together with ln (Educ) , and ln (Innov) with ln (ICT) because of their respective high correlations (0.862 and 0.766). The level of multicollinearity between the explanatory variables remains generally high and represents an issue that will be considered in the following section.

Furthermore, it is important to avoid encountering spurious relationships with the presence of variables integrated of order 2. To test for stationarity, a variety of unit root tests are performed. The Levin, Lin and Chu (2002) test assumes common unit root, and the Im, Pesaran and Shin (2003) and the Fisher ADF (Choi 2001) tests assume individual unit root. The findings are reported in Table A4. The results from the three unit root tests agree on the degrees of stationarity of ln (*TFP*) , which is of order 0, and *ln (FDI) , ln (Import) , ln (Educ) and ln (Institu)* which are of order 1. The remaining variables *ln (Innov)* and *ln (ICT)* are of order 0 under the common unit root test and of order 1 under

the individual unit root tests. Nevertheless, those findings reveal that, independent of which test is being used, no variable is integrated of order 2. Hence, this confirms that the panel ARDL model can be employed.

After the confirmation of mixed stationarity status, the Pedroni (1999, 2004) and Kao (1999) panel cointegration tests are performed to examine the existence of a long-run relationship between knowledge components and *TFP*. Pedroni considers both pooled-within dimension tests and group-mean-between dimension tests. The results are depicted in Table A5 and A6. On the one hand, 2 of the 11 Pedroni statistics reject the null hypothesis of no cointegration. On the other hand, the Kao residual panel cointegration test reveals that the null hypothesis of no cointegration is rejected at the 1% significance level. I conclude that there exists a long-run relationship among the variables.

4 Empirical Results

After a series of trial regressions, it appears that the pillars of the KE are relatively fragile, in the sense that the signs and/or statistical significance of the coefficients vary between different model specifications (i.e. the combination of variables included). Because all components of the KE are interconnected, the issues of multicollinearity and omitted variables weaken the robustness of the estimated coefficients. The Extreme Bounds Analysis (EBA) is a helpful statistical exercise developed by Learner (1983) that investigates the robustness of explanatory variables for all possible combinations. However, I do not possess the statistical package that is required to implement an EBA. Alternatively, I run a large series of regressions for all possible combinations of explanatory variables, taking into consideration multicollinearity. Table A7 is a summary of the signs and significance levels of the long-run ARDL-PMG estimators for such combinations.

Firstly, in panel (I), I regress *TFP* on each explanatory

variable individually. Each variable appears as a significant driver of *TFP* but ln (Import) . Moreover, it has been recognized earlier that ln (Import) is highly correlated with ln (Educ) . Therefore, the temptation is to exclude ln (Import) . Because the variable *ln (FDI)* is already an indicator of knowledge spillovers, this exclusion does not conceal the representation of foreign knowledge spillovers in the model. Secondly, it can be inferred from the multiple model specifications in panel (II) that all the variables are, indeed, relatively sensitive. The least fragile variables, whose coefficients remain significant and of the same sign most of the time irrespective of the model specification are *ln (Innov) , ln (Educ) , ln (ICT)* and *ln (Institu)*. The first three variables remain significantly positive and ln (Institu) significantly negative, most of the time. Building on these sensitivity results, I select the most appropriate model specifications, i.e., the regressions which reproduce those recurrent relationships. These models are illustrated in Table 1. In order to avoid multicollinearity, only three pillars of KE are included at a time. Column (a) cumulatively illustrates the innovative system (internal innovation), the educational system, and the institutional regime. In column (b), the innovative system is indicated by the external innovation. Column (c) shows the educational and institutional systems together with the ICT pillar.

The first panel presents the short-run ARDL-PMG estimates. The ECT, which indicates the speed of adjustment from short-run disequilibrium to long-run equilibrium, is expected to have a negative sign. In the case at hand, the ECT is always negative and of magnitude between -0.021 and -0.066. This suggests that the deviation of variables from the short-run to the long-run equilibrium is adjusted and corrected by 0.021% to 0.066% annually. However, the estimated coefficients are not statistically significant. The first and second lagged dependent variables

5 - The optimal lag length of the panel ARDL (3, 2, 2, 2, 2, 2, 2) is selected according to the Akaike Information Criterion (AIC).

Table 1. Panel ARDL-PMG Estimates

	Panel 1: SHORT-RUN ESTIMATES		
	(a)	(b)	(c)
ECT	-0.021	-0.042	-0.066
	(-0.177)	(-0.522)	(-0.569)
Δln(TFP)(-1)	0.074	0.312	0.338**
	(0.835)	(2.190)	(2.476)
Δln(TFP)(-2)	-0.135	-0.207**	-0.238
	(-1.289)	(-2.527)	(-1.698)
Δln(Innov)	-0.013		
	(-0.021)		
Δln(Innov)(-1)	1.032*		
	(1.967)		
Δ Ln(FDI)		0.100***	
		(0.470)	
Δln(FDI)(-1)		-0.053	
		(-0.997)	
Δln(Educ)	-2.760	-2.020	-0.259
	(-0.302)	(-0.208)	(0.026)
Δln(Educ)(-1)	-10.011**	-3.362	2.415
	(-2.689)	(-1.271)	(0.513)
Δln(ICT)			0.037
			(0.291)
Δln(ICT)(-1)			0.021
			(0.228)
Δln(Institu)	0.935	0.938	0.453
	(0.895)	(1.132)	(0.533)
Δln(Institu)(-1)	0.498	0.631	0.200
	(0.958)	(0.904)	(0.673)
C	0.064	0.740	0.304
	(0.050)	(0.580)	(0.667)
	Panel 2: LONG-RUN ESTIMATES		
	(a)	(b)	(c)
ln(Innov)	0.323**		
	(2.012)		
ln(FDI)		-0.132	
		(-1.317)	
ln(Educ)	4.669**	3.454**	1.983
	(2.164)	(2.497)	(1.506)
ln(ICT)			0.490***
			(2.642)
ln(Institu)	-3.634***	-3.072**	-1.471**
	(-5.014)	(-2.233)	(-1.996)

Source: Author's calculations. Note: t-statistics in parentheses. *p<0.1, **p<0.05, ***p<0.01

have positive and negative coefficients respectively. This suggests that *TFP* in the current year depends positively on the change in *TFP* from the previous year but negatively from that of two years prior. Nevertheless, these estimates are only statistically significant at the 5% significance level once that domestic innovation is not controlled for.

The lagged indicator of domestic innovation seems to have a positive impact on *TFP* in the short run at the 10% significance level. This suggests that a 1% increase in the stock of resident patent applications in the previous year increases *TFP* by 1.032% in the current year, ceteris paribus. The estimate for knowledge transmission in the current year appears to be significantly positive. A 1% increase in the inward stock of FDI in the current year increases *TFP* directly by 0.1% in the same year, ceteris paribus. Next, a surprising observation concerns the coefficient of education and skills in the short run. According to the estimates, the human capital index negatively influences *TFP* in the short run. Nevertheless, only lagged estimate of the model including domestic innovation is significant at the 5% significance level. Furthermore, the two remaining pillars, indicated by ICT and institutional quality, seem to positively influence *TFP* in the short run. However, the estimated coefficients are not significant. In conclusion, the short-run estimates of the KE variables are mainly insignificant. This signifies that the four pillars which promote the accumulation, creation, and efficient use of knowledge do not seem to have an immediate impact on *TFP* levels in the six sampled countries, probably because of adjustment efforts.

The existence of a long-run relationship among variables that was ascertained in Pedroni (1999, 2004) and Kao (1999) cointegration tests in the previous section is again inferred from the statistical significance of the long-run coefficients in panel 2.

With respect to the effects of KE components on *TFP* levels, the long-run PMG estimates do not always support

the short-run findings. The first variable, representing domestic innovation, appears as a significant driver of *TFP* enhancements at the 5% significance level. However, the magnitude of the coefficient remains relatively small. In fact, a 1% increase in the stock of resident patent applications leads to a 0.323% increase in *TFP* in the long run, ceteris paribus. Nevertheless, it is important to recognize that domestic innovation is proxied by the stock of resident patent applications, instead of grants. Hence, only some of these registered patents are granted and can be exploited. Thus, this proxy represents the outcome of R&D activities – the production of ideas – but not the degree to which these ideas or innovations are exploited in the production process or adopted by households. This measurement may affect the significance and magnitude of the estimate. On the other hand, the coefficient of the indicator of foreign knowledge spillovers does not have the expected positive sign and is insignificant. As inferred from Table A7, ln (FDI) is the most fragile variable and according to Learner (1985:1), "a fragile inference is not worth taking seriously."

The second pillar of the KE, education and skills, represents a substantial driver of *TFP* long-run improvements, contradicting the previous short-run findings. According to the long-run estimates, a 1% increase in the index of human capital leads to about 4% increase in *TFP*, ceteris paribus. Those results satisfy the earlier stated hypothesis.

Similarly, the long-run estimate of ICT is positive and significant. This suggests that in the six sampled countries, the more individuals that use the Internet, the higher the *TFP* levels in the long run.

The indicator of economic and institutional regime has a significantly negative impact on *TFP* levels in the long run. In other words, the social infrastructures in the six leading Asian knowledge-based economies have contributed negatively to *TFP* levels. Despite the fact that these results contradict the hypothesis of positive influence, they seem

consistent with most of the model specifications in Table A7. This contradicts the findings of Rodrik (1997), who emphasized the importance of institutional quality in implementing the right policies fostering *TFP* growth in East Asia.

5 Policy Implications

It is hazardous to draw conclusions from data that are relatively fragile. Nevertheless, a few policy suggestions may be inferred from the least fragile recurrent relationships that have appeared at this stage. It is clear from the above exercise that measures are expected to differ depending on whether the objective is to foster *TFP* improvements in the short run or in the long run. In the case at hand, on the topic of sustainable productivity-led growth, the target should be on the long-run impact of the KE pillars. The long-run estimates suggest that there are three key areas in which knowledge can foster *TFP* enhancements in the six sampled countries.

The long-run elasticity of *TFP* with respect to education and skills is substantial. This suggests a human capital-led *TFP* growth in the selected economies, supporting the new growth theories (Romer, 1986; Lucas, 1988). The quality of education has been the priority of the East Asian welfarist system (Gough, 2004a; Aspalter, 2011). Those economies have taken into consideration the fundamental changes of the century and have shifted the focus of their educational system accordingly. Their ideology of "learning to learn" revolves around three emerging needs: (1) learning and innovation, (2) digital literacy, and (3) social and emotional skills (Kattan and Bend, 2018). This focus on the accumulation of appropriate knowledge and competencies needed in knowledge-based economies may be the reason why human capital has been a significant driver of *TFP* improvements. This modern approach to learning should be maintained in the future. However, one issue that policymakers should consider while increasing

human capital is the potential issue of "brain drain." In East Asia, an increasing number of educated and high-skilled individuals seek better career opportunities abroad (Yap, 2017).

The promotion of a dynamic innovation system is also needed to achieve a successful growth of *TFP* in the long run. Although the elasticities of *TFP* to foreign knowledge appear larger in East Asia than in OECD countries in previous research, knowledge transfers through the FDI channel is highly fragile and mostly insignificant in this study. However, *TFP* growth can be relatively more robustly attributed to domestic innovation. In order to build a stronger "Innovative Asia," as claimed by the ADB, East Asian economies should ensure the protection of intellectual property rights and invest in the R&D sectors, from universities to research institutes.

In this study, the productivity paradox has been defied. In fact, the availability of ICT appears to be a third means of enhancement of *TFP* levels. The inclusion of technology in the innovative and educational systems and in the business environment should remain a policy priority in East Asia. The promotion of the digitally-enabled economy was the subject matter of the 13th East Asian Summit. The ASEAN ICT Masterplan (AIM) 2020 then discussed is an example of an appropriate measure promoting affordable access to digital technologies and aiming towards the decline of the digital divide in East Asia (ASEAN 2016).

According to the long-run estimators, institutional quality has had a significantly negative effect on *TFP* in the long run in the six sampled East Asian economies. These unexpected results contradict with Rodrik's (1997) findings and with my previously stated hypothesis. Nevertheless, it would be a wrong inference to draw fast conclusions from these empirical results with respect to the influence of the fourth pillar. On the one hand, the different indices representing the quality of institutional regimes as retrieved from the PRS Group might have been improperly presented

or measured. On the other hand, the quality of institutions in establishing policies targeting knowledge accumulation is not only captured by the index used in the exercise. In fact, it is also partly revealed through the implementation of appropriate programs for the three other pillars. For instance, governments that have carried out efficient and effective measures in the objective to foster capacity-building education is captured in the estimate coefficients of Educ. The adoption of the right policies at the right time may have been a crucial driver of *TFP*, as suggested by Nelson and Pack (1999).

Although those findings are applicable to the sampled group consisting of Hong Kong, Japan, Singapore, Republic of Korea, Malaysia, and Thailand, the ADB advised the rest of Asia in its transition towards knowledge-based economies to take the experiences of those leaders as examples. Thus, not only do the following policy suggestions apply to the Asian leaders, but they may also guide policymakers in developing Asia to take the right paths to move faster towards the technology frontier. These KE factors should not be seen in isolation, however. Depending on the country's initial level of development, a set of essential factors (i.e., sanitation), may be considered first before shaping the country's status towards knowledge-based players.

6 Limitations and Future Research

The results of this exercise need to be interpreted with care since the dependent variable is *TFP*, one of the most ambiguous concepts in economics. *TFP* is the residual that measures what cannot be explained by the observable input factors. Abramovitz (1956:11) goes so far as to call it the "measure of our ignorance." The controversy around the concept, measurement, and interpretation of *TFP* has been the subject of a large literature. The concept of *TFP* depends critically on its arbitrary definition and is based on strong assumptions. Hence, there is ample scope for mismeasurement and ambiguity. In this study, the calculation of *TFP* was founded on bold assumptions,

following Solow (1957). Eventually, the aggregate production function was assumed to be in the Cobb Douglas form, displaying constant returns to scale. Future studies should reconsider the use of more sophisticated methods of measurements of *TFP* and verify if the results described in this paper are consistent across different measurement approaches. For instance, non-parametric approaches do not rely on the predetermined specification of any functional form. For instance, the Törnqvist index is a non-frontier approach which computes the weighted differences in the growth rates of outputs and inputs, as seen in Diewert (1976). Another possibility is the Malmquist index which is a frontier approach that calculates the ratio of the distances of each data point relative to a common frontier (Caves et al., 1982).

Furthermore, *TFP* is largely sensitive to the measurement of input factors. Because of a lack of available data, the capital utilization was not considered in this study. However, according to Burnside et al. (1995) and Basu (1996), failing to account for capital utilization rates may lead to under- or over-estimation of *TFP* figures. Therefore, correcting for cyclical variations in capital services is necessary to account for the true variation of *TFP* levels over time. Although these possible future studies may help to relax some of my assumptions, one cannot escape making some assumptions.

Additionally, this study analyzes aggregate *TFP* and, hence, does not discriminate between different sectors of production. Nevertheless, the four pillars of the KE might not have similar impact on *TFP* levels across all industries. Future micro-level research would allow a more precise analysis of the knowledge-*TFP* relationship in different sectors of production or firms. It would allow the formulation of more sensitive policies in the future. Moreover, the large inherent variation in firm-level data should neutralize errors and ambiguities stemming from the measurement of *TFP* (Syverson, 2011).

The accuracy of empirical results is additionally hindered

by the scarcity and inadequacy of data. The series compiled were in inadequate form and, hence, subject to multiple manipulations. For instance, the calculation of knowledge stocks was based on strong assumptions concerning the depreciation rate of patents. Moreover, the scarcity of data series highly constrained the selection of proxies for the indicator of each of the four pillars. Consequently, it is not clear whether the results of Table 1 represent the real phenomena that were meant to be measured or are just another reflection of insufficient data. For robustness matters, future studies should envisage the analysis of different proxies and indicators. For instance, the innovation system could be proxied by domestic R&D expenditures, whereas foreign R&D expenditures could indicate knowledge spillovers. Moreover, education and skills could be proxied by the measure of cognitive skills constructed by Hanushek and Kimko (2000). Additionally, the Classification of the Functions of Government (COFOG) could be considered to investigate into gross expenditures into knowledge domains, as an indicator of government interventionism rather than of institutional quality. Similarly, future research may consider some of the KE normalized proxies compiled in the Knowledge Assessment Methodology (KAM) (Chen and Dahlman, 2006). Interaction variables could be examined too. More specifically, future research subject should analyze the effect of interacting education with the indicators of each of the other three pillars.

6 Conclusion

In light of the major role of TFP in raising living standards and fostering sustainable economic prosperity, new growth theories have acknowledged the role of knowledge as the heart of economic development (Romer, 1986; Lucas, 1988). In 1999, the World Bank published a symbolic report advancing strategies for building KEs and underlined four key ingredients for a successful transition: (1) the innovation system, (2) education and skills, (3) ICT infrastructure, and

(4) the economic and institutional regime. The progression towards knowledge-based economies has been disparate in Asia and only a handful of Asian economies have led their transition. This research work fills the gaps in the existing literature as it aims at analyzing the relationship between the four knowledge pillars and TFP for six successful Asian knowledge-based economies. The ARDL-PMG approach by Pesaran et al. (1999) is applied on the period spanning between 1996 and 2017. Nevertheless, the statistical illustration of the intimate relationship between knowledge and TFP is laborious and the results lack robustness. Nevertheless, based on the least fragile variables and the most recurrent relationships, the study concludes that three knowledge pillars have enhanced TFP in the six sampled economies: domestic innovation, the level of education and skills, and the access to ICT. 25 years after the alarming predictions of Krugman (1994) concerning the future of East Asia, my results suggest that, thanks to the contribution of those three KE pillars, a persistent positive growth path will be sustained. Moreover, developing Asian economies are advised to shape their own priorities in the lens of these findings. Taking example of the experience of the Asian leaders may help them to pursue their transition and improve their national productivity. I believe that this study can guide the transformation of Asia into an advanced knowledge-based player.

7 Bibliography

Abramovitz, Moses. 1956. "Resource and Output Trends in the United States since 1870." American Economic Review 5-23.

Ahmed, Elsadig Musa. 2017. "ICT and Human Capital Spillover Effects in Achieving Sustainable East Asian Knowledge-Based Economies." Journal of the Knowledge Economy 8: 1086–1112.

Ang, James B., and Jakob B. Madsen. 2011. "Can Second-

Generation Endogenous Growth Models Explain The Productivity Trends and Knowledge Production in the Asian Miracle Economies?" The Review of Economics and Statistics 93 (4): 1360-1373.

ASEAN. 2016. The ASEAN ICT Masterplan 2020. Jakarta: The Association of Southeast Asian Nations.
Asian Development Bank. 2014. Innovative Asia: Advancing the Knowledge-based Economy. Mandaluyong City: Asian Development Bank.

Asian Development Bank. 2007. Moving Toward Knowledge-Based Economies: Asian Experiences. Regional and Sustainable Development Department, Metro Manila: Asian Development Bank, 1-187.

Aspalter, Christian. 2011. "The Development of Ideal-typical Welfare Regime Theory." International Social Work 54 (6): 735-750.

Barro, Robert J., and Jong-Wha Lee. 2013. "A new data set of educational attainment in the world, 1950-2010." Journal of Development Economics (104): 184–198.

Basu, Susanto. 1996. "Procyclical Productivity: Increasing Returns or Cyclical Utilization." Quarterly Journal of Economics 111: 719-751.

Birdsall, Nancy M., Jose Edgardo L. Campos, Chang-Shik Kim, W. Max Corden, Lawrence [editor] MacDonald, Howard Pack, John Page, Richard Sabor, and Joseph E. Stiglitz. 1993. The East Asian miracle : Economic Growth and Public Policy. New York: World Bank Oxford University Press.

Brynjolfsson, Erik. 1993. "The Productivity Paradox of Information Technology." Communications of the ACM 36 (12): 66-77.

Burnside, Craig, Martin Eichenbaum, and Sergio T. Rebelo.

1995. "Capital Utilization and Returns to Scale." NBER Macroeconomics Annual 1995 67-110.

Caves, Douglas W., Laurits R. Christensen, and Walter Diewert. 1982. "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity." Econometrica 50 (6): 1393-1414.

Chen, Derek H.C., and Carl J. Dahlman. 2006. The Knowledge Economy, the KAM Methodology and World Bank Operations (English). Washington DC: World Bank.

Choi, In. 2001. "Unit Root Tests for Panel Data." Journal of International Money and Finance 20 (2): 249-272.

Coe, David T., and Elhanan Helpman. 1995. "International R&D spillovers and institutions." European Economic Review 39: 859–887.

Cohen, Daniel, and Laura Leker. 2014. Health and Education: Another Look with the Proper Data. Paris School of Economics, Paris: mimeo.

Cohen, Daniel, and Marcelo Soto. 2007. "Growth and Human Capital: Good data, Good Results." Journal of Economic Growth 12 (1): 51-76.

Diewert, Walter Erwin. 1976. "Exact and Superlative Index Number." Journal of Econometrics 4 (2): 115-145.

Evenson, Robert E., and Lakhwinder Singh. 1997. Economic Growth, International Technological Spillovers and Public Policy: Theory and Empirical Evidence from Asia. Working Paper 777, Economic Growth Center, Yale University, New Haven: Econstor.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." American Economic Review 105 (10): 3150-3182.
Gough, Ian. 2004a. "East Asia: The Limits of Productivist

Regimes." In Insecurity and Welfare Regimes in Asia, Africa, and Latin America: Social Policy in Developmental Contexts, by Ian Gough, Geoffrey D Wood, Armando Barrientos, Philippa Bevan, Peter Davis and Graham Room, 169-201. Cambridge: Cambridge University Press.

Griliches, Zvi. 1980. "R&D and the Productivity Slowdown." American Economic Review 343-348.

Hall, Bronwyn H., Zvi Griliches, and Jerry A. Hausman. 1986. "Patents and R&D: Is There a Lag?" International Economic Review 27 (2): 265-283.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" The Quarterly Journal of Economics (Oxford University Press) 114 (1): 83-116.

Hanushek, Eric A., and Dennis, D. Kimko. 2000. ""Schooling, Labour Force Quality, and the Growth of Nations." American Economic Review 90: 1184-1208.

Hicks, John Richard. 1966. The Theory of Wages. New York: MacMillan.

Im, Kyung So, Harlem M. Pesaran, and Yongcheol Shin. 2003. "Testing for Unit Roots in Heterogeneous Panels." Journal of Econometrics 115 (1): 53-74.

Isaksson, Anders. 2007. Determinants of Total Factor Productivity: a Literature Review. Staff Working Paper 02/2007, Research and Statistics Branch, United Nations Industrial Development Organization, Vienna: UNIDO Publications.

Kao, Chihwa. 1999. "Spurious Regression and Residual-based Tests for Cointegration in Panel Data." Journal of Econometrics 90: 1-44.

Kattan, Raja Bentaouet, and May Bend. 2018. "Educating

for the future: The case of East Asia." World Bank. 12 September. https://blogs.worldbank.org/education/educating-future-case-east-asia.

Kim, Jong-Il, and Lawrence J Lau. 1994. "The Sources of Economic Growth of the East Asian Newly Industrialized Countries." Journal of the Japanese and International Economies 8 (3): 235-271.

Knack, Stephen, and Philip Keefer. 1995. "Institutions and Economic Performance: Cross-country Tests Using Alternative Institutional Measures." Economics and Politics 7 (3): 207-227.

Knowles, Stephen, and Arlene Garces-Ozanne. 2003. "Government Intervention and Economic Performance in East Asia." Economic Development and Cultural Change 51 (2): 451-477.

Kraemer, Kenneth L., and Jason Dedrick. 1994. "Payoffs from Investment in Information Technology: Lessons from the Asia-Pacific Region." World Development 22 (12): 1921-1931.

Krugman, Paul R. 1994. "The Myth of Asia's Miracle." Foreign Affairs, 73 (6): 62-78.

Leamer, Edward E. 1985. "Sensitivity Analyses Would Help." American Economic Review 75 (3): 308-313.

Leamer, Edward E. 1983. "Let's Take the Con Out of Econometrics." American Economic Review 73 (1): 31-43.
Levin, Andrew, Chien-Fu Lin, and Chia-Shang James Chu. 2002. "Unit root tests in panel data: asymptotic and finite-sample properties." Journal of Econometrics 108 (1): 1-24.

Lucas, Robert E. 1988. "On the Mechanics of Economic Development." Journal of Monetary Economics (Elsevier Science Publishers B.V.) 22: 3-42.

Madden, Gary G, Scott J Savage, and Paul Bloxham. 2001. "Asian and OECD international R&D spillovers." Applied Economics Letters 431-435.

Mahmood, Ammara, and Talat Afza. 2008. "Total Factor Productivity Growth in East Asia: A Two Pronged Approach." European Journal of Economics, Finance and Administrative Sciences 14 (14): 93-113.

Nelson, R. R., and H. Pack. 1999. "The Asian Miracle and Modern Growth Theory." The Economic Journal 109 (457): 416-436.

Nelson, Richard R., and Edmund S. Phelps. 1966. "Investment in Humans, Technological Diffusion, and Economic Growth." American Economic Association (American Economic Association) 56 (1/2): 69-75.

Okabe, Misa. 2002. "International R&D Spillovers and Trade Expansion Evidence from East Asian Economies." ASEAN Economic Bulletin (Institute of Southeast Asian Studies) 19 (2): 141-154.

Olson, Mancur Jr. 1996. "Distinguished Lecture on Economics in Government: Big Bills Left on the Sidewalk: Why Some Nations are Rich, and Others Poor." The Journal of Economic Perspectives 10 (2): 3-24.

Park, Donghyun, and Jungsoo Park. 2010. Future, Drivers of Developing Asia's Growth: Past and Future. ADB Economics Working Paper Series No. 235, Manila: Asian Development Bank.

Park, Gwangman, Juneseuk Shin, and Yongtae Park. 2006. "Measurement of Depreciation Rate of Technological Knowledge: Technology Cycle Time Approach." Journal of Scientific & Industrial Research 65: 121-127.

Pedroni, Peter. 1999. Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors.

Department of Economics Working Papers, Department of Economics, Williams College, Massachusetts: Oxford Bulletin of Economics and Statistics.

Pedroni, Peter. 2004. Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis. Department of Economics Working Papers, Department of Economics, Williams College, Massachusetts: Oxford Bulletin of Economics and Statistics.

Pesaran, M. Hashem, Yongcheol Shin, and Ron P. Smith. 1999. "Pooled Mean Group Estimation of Dynamic Heterogeneous Panels." Journal of the American Statistical Association 94 (446): 621-634.

PRS Group. 2019. "CountryData Online (CDO)." Political Risk Services. 25 January. https://epub.prsgroup.com/index.php/country-database/country-data.

Psacharopoulos, George. 1994. "Returns to Investment in Education: A Global Update." World Development 22 (9): 1325–1343.

Rodrik, Dani. 1997. TFPG Controversies, Institutions, and Economic Performance in East Asia. NBER Working Paper 5914, Cambridge: The National Bureau of Economic Research, 1-43.

Romer, Paul M. 1990. "Endogenous Technological Change." Journal of Political Economy (The University of Chicago Press) 98 (5): S71-S102.

Romer, Paul M. 1986. "Increasing Returns and Long-Run Growth." The Journal of Political Economy (The University of Chicago Press) 94 (5): 1002-1037.

Solow, Robert M. 1956. "A Contribution to the Theory of Economic Growth." The Quarterly Journal of Economics (Oxford University Press) 70 (1): 65-94.

Solow, Robert M. 1957. “Technical Change and the Aggregate Production Function.” The Review of Economics and Statistics (The MIT Press) 39 (3): 312-320.

Solow, Robert. 1987. “We'd better watch out.” New York Times Book Review, 12 July: 36.

Syverson, Chad. 2011. “What Determines Productivity?” Journal of Economic Literature (American Economic Association) 49 (2): 326-365.

Thomas, Vinod, and Yan Wang. 1996. “Distortions, Interventions, and Productivity Growth: Is East Asia Different?” Economic Development & Cultural Change 44 (2): 265-288.

UNCTAD. 2019. World Investment Report 2019. United Nations Conference on Trade and Development, Geneva: United Nations Publications . https://unctad.org/en/Pages/DIAE/World%20Investment%20Report/Annex-Tables.aspx.

WIPO. 2018. “WIPO Statistics Database.” World Intellectual Property Organization. 1 December. https://www.wipo.int/pct/en/.

World Bank. 2019. “World Development Indicators.” World Bank Open Data. 25 June. https://data.worldbank.org/.

World Bank. 1999. World Development Report 1998/1999 : Knowledge for Development. The International Bank for Reconstruction and Development , Washington: Oxford University Press, 266.

Yap, Karl Lester M. 2017. “Brain Drain from Southeast Asia Poses Obstacle to Growth.” Bloomberg. 5 June. https://www.bloomberg.com/news/articles/2017-06-05/brain-drain-from-southeast-asia-poses-obstacle-to-growth.

Young, Alwyn. 1994. “Lessons from East Asian NICs: a

Contrarian View." European Economic Review 38 (3-4): 964-973.

Yuhong, Du, T. Sarmidi, and Abu Hassan Shaari Mohd Nor. 2017. "Nonlinear effects of human capital on TFP: Evidence from eight ASEAN countries." Journal of Economic Cooperation and Development 38 (4): 123-153.

Zachariadis, Marcos. 2004. "R&D induced growth in the OECD?" Review of Development Economics 8: 423-439.

List of Tables and Figures

8 Appendix

Table A1

Solow Residuals.

	Hong Kong	Japan	Singapore	Rep. of Korea	Malaysia	Thailand
1995	44,45	284,90	37,20	124,36	8,09	9,30
1996	50,75	265,08	37,48	137,31	7,97	10,10
1997	53,74	261,00	36,91	129,33	7,87	9,58
1998	59,01	270,96	38,89	93,89	7,18	9,87
1999	69,20	265,57	44,72	81,44	7,48	10,69
2000	73,02	248,52	44,89	77,88	7,83	10,37
2001	75,28	235,53	52,55	78,56	7,69	10,94
2002	78,67	220,24	59,69	81,81	7,88	11,64
2003	80,31	197,81	52,98	83,95	8,11	12,43
2004	78,16	180,58	42,65	87,82	8,48	13,22
2005	69,96	171,48	33,98	92,67	8,67	13,72
2006	67,56	171,58	33,93	102,66	8,89	13,85
2007	71,21	169,65	35,43	106,25	9,89	13,87
2008	72,42	179,82	40,04	104,26	12,64	14,37
2009	79,67	200,67	47,41	99,37	15,20	15,13
2010	99,65	203,33	49,54	93,21	16,18	15,56
2011	105,48	202,45	45,63	90,10	15,99	14,37
2012	108,52	216,12	45,41	96,24	17,25	15,22
2013	111,30	206,37	45,78	102,71	18,91	15,53
2014	113,00	198,69	46,06	109,25	20,58	15,71
2015	111,03	192,84	45,91	114,80	21,37	16,15
2016	110,27	193,27	46,04	118,15	21,46	16,49
2017	114,69	204,51	47,13	116,96	21,77	16,94

Source: Author's calculations.

Table A2

Variables definition and descriptive statistics.

Variable name	Definition	Source	Obs.	Mean	Median	Min.	Max.	Std. Dev.
TFP	Total Factor Productivity (constant 2011 mil.US$)	Penn World Table 9.1	138	78.0	52.8	7.2	284.9	72.0
Innov	Stock of resident patent applications	World Intellectual Property Organisation	138	516636.5	5071.3	113.0	2689253.5	926320.1
FDI	Inward Stock of Foreign Direct Investments (constant 2011 mil.US$)	United Nations Conference on Trade and Development	138	278984.4	133265.2	19439.3	1653157.2	364099.1
Import	Total imports (constant 2011 mil.US$)	World Bank Data	138	334705.0	244542.6	67012.35	1005842.1	208644.5
Educ	Human Capital Index	Penn World Table 9.1	138	2.99	3.00	2.14	3.97	0.42
ICT	Individuals using the Internet (per thousand indiv.)	World Bank Data	138	76.9	36.8	1.8	351.9	88.5
Institu	Weighted Average of Six Components of the Political Risk Rating Index	International Country Risk Guide	138	66.7	68.4	46.4	79.7	8.4

Source: Author's calculations.

Table A3

Correlation Matrix.

	ln(TFP)	ln(Innov)	ln(FDI)	ln(Import)	ln(Educ)	ln(ICT)	ln(Institu)
ln(TFP)	1.000						
ln(Innov)	0.659	1.000					
ln(FDI)	0.230	-0.318	1.000				
ln(Import)	0.766	0.598	0.505	1.000			
ln(Educ)	0.797	0.690	0.347	0.862	1.000		
ln(ICT)	0.422	0.766	-0.034	0.591	0.630	1.000	
ln(Institu)	0.695	0.419	0.176	0.567	0.595	0.061	1.000

Source: Author's calculations.

Table A4

Panel Unit Root Tests Results.

Variable	Levin, Lin & Chu		Im, Pesaran and Shin		ADF-Fisher	
	Null: Unit root (common unit root process)		Null: Unit root (individual unit root process)			
	Stat.	Prob.	Stat.	Prob.	Stat.	Prob.
ln(TFP)	-4.756***	0.0000	-4.051***	0.0000	40.894***	0.0001
Δln(TFP)	-5.458***	0.0000	-4.123***	0.0000	38.044***	0.0002
ln(Innov)	-3.013**	0.0013	-1.401*	0.0806	22.844**	0.0291
Δln(Innov)	-2.125**	0.0168	-3.528***	0.0002	37.803***	0.0002
ln(FDI)	-1.580*	0.0570	-0.148	0.4410	12.075	0.4397
Δln(FDI)	-10.583***	0.0000	-9.767***	0.0000	83.022***	0.0000
ln(Import)	0.572	0.7162	0.148	0.5588	9.518	0.6582
Δln(Import)	-7.721***	0.0000	-5.649***	0.0000	47.948***	0.0000
ln(Educ)	-0.008	0.4969	1.126	0.8699	10.170	0.6011
Δln(Educ)	-1.965**	0.0247	-1.669**	0.0475	19.8151*	0.0707
ln(ICT)	-11.891***	0.0000	-15.435***	0.0000	140.436***	0.0000
Δln(ICT)	-0.785	0.2161	-3.903***	0.0000	35.819***	0.0003
ln(Institu)	-2.634***	0.0042	-2.434***	0.0075	24.569**	0.0170
Δln(Institu)	-7.554***	0.0000	-5.804***	0.0000	49.994***	0.0000

Source: Author's calculations. Note: Specification of intercept with a trend. *p<0.1, **p<0.05, ***p<0.01.

Table A5

Pedroni Cointegration Test Results.

Tests	Statistic	Prob.	Weighted Statistic	Prob.
Within-Dimension				
Panel v-Statistic	0.328	0.3715	-0.980	0.8364
Panel rho-Statistic	2.782	0.9973	2.683	0.9963
Panel PP-Statistic	1.679	0.9534	0.799	0.7878
Panel ADF-Statistic	-1.266	0.1028	-1.848**	0.0323
Between-Dimension				
Panel rho-Statistic	3.785	0.9999		
Panel PP-Statistic	1.192	0.8833		
Panel ADF-Statistic	-1.740**	0.0409		

Source: Author's calculations. Note: Specification of intercept with a trend. *p<0.1, **p<0.05, ***p<0.01.

Table A6

Kao Residual Panel Cointegration Test Results.

Test	t-statistics	Prob.
ADF	-2.400***	0.0082

Source: Author's calculations. Note: Specification of intercept with a trend. *p<0.1, **p<0.05, ***p<0.01.

The Impact of the Introduction of the UK National Living Wage on the Employment Probabilities of Low-Wage Workers

Brian O'Connor
Univeristy of Nottingham
Economics

Abstract

This This paper adopts a difference-in-difference methodology first employed by Linnerman (1982) to determine how the 2016 UK National Living Wage has affected subsequent employment probabilities of those with low wages. Longitudinal data has been sourced from four consecutive Labour Force Surveys straddling the implementation date (1st April 2016) of the new minima in order to determine this effect. Estimates suggest there are negative effects on employment for those on low wages that are statistically significant from zero and increasing with the duration of time analysed. Regional tests present evidence that regions of medium incidence of low pay are the worst affected areas, while sex tests conclude men are more adversely affected than women, although these results lack statistical significance.

"The central challenge facing policy makers when introducing minimum wage legislation is to raise the pay of low paid individuals without harming their employment prospects." – (R. Dickens, R. Riley and D. Wilkinson, 2015)

Table of Contents

1 Introduction

On the 1st of April 2016, the National Living Wage (*NLW*) was introduced in the UK. The aim of the policy was to raise wages for more experienced workers by introducing a premium for those aged 25 and over. The premium took the form of an additional 50p per hour bringing the total National Minimum Wage (*NMW*) to £7.20. This change represents the largest nominal increase in the UK minima since the introduction of the *NMW* on 1st April 1999.

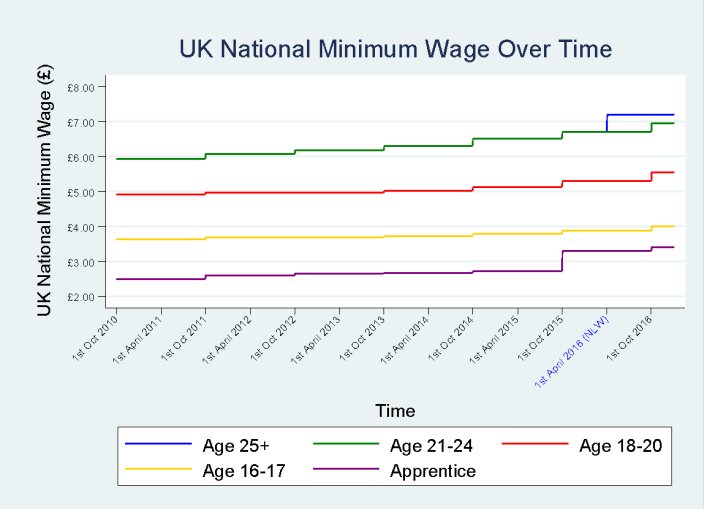


Figure 1 - UK National Minimum Wage over Time

As noted by Dickens et al., (2015) "the central challenge facing policy makers when introducing minimum wage legislation is to raise the pay of low paid individuals without harming their employment prospects." The *NLW* was introduced to tackle poverty, however, if its implementation has led to individuals losing their jobs then the policy may be counterproductive to its aims. This paper seeks to determine the effect the *NLW* has on employment probabilities of low paid workers in the U.K., utilizing a standard difference-in-difference methodology first adopted by Linnerman (1982). Section

2 of this paper gives an overview of the vast literature that has been published on the employment effects of legislative minima. Sections 3 and 4 explain the estimation strategy and data that will be used in order to conduct analyses on the effect of *NLW*. Results for the basic specification and checks of robustness are presented in sections 5 and 6 respectively. Section 7 provides concluding remarks of the analysis.

2 Literature Review

2.1 Economic Theory

Employment implications of national minima have long been a pivotal source of debate within labour economic literature. Traditional theory suggests that a binding minimum wage in a single competitive labour market with homogenous workers will lead to a reduction in employment, given that the minimum is set above equilibrium (Figure 2).

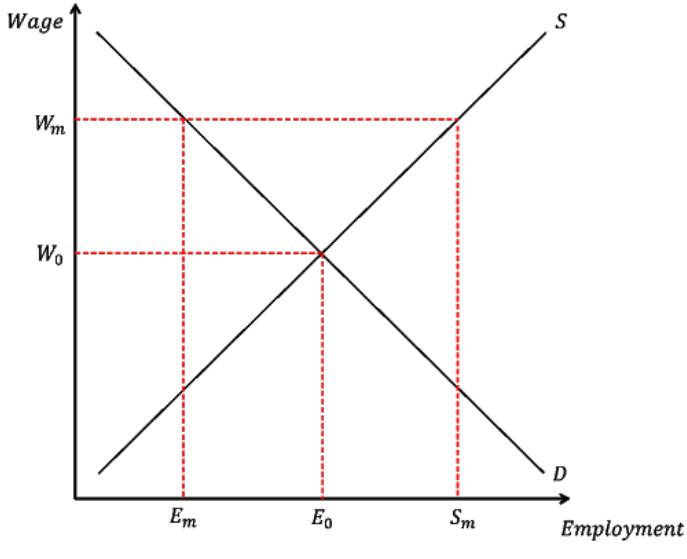


Figure 2 - Traditional Economic Theory Graph

In this model "the proportional reduction in employment (lnEm-lnE0) is equal to the proportional wage increase (lnWm-lnW0) times the elasticity of demand" (Brown et al., 1982). Under the assumptions of traditional theory, one would assume that the implementation of a *NLW* would reduce employment given the minimum is set above the equilibrium level. However, traditional theory is starkly contradicted by models that assume employers have labour market power—a monopsony—and in this case show that it is possible for the minimum wage to increase employment (West and McKee, 1980). Opposition also arises from shock effect, efficiency wages, and job search models (Jardim et al., 2017).

2.2 Empirical Studies

A plethora of literature has been written on the empirical effects of pay floors using a number of different data sources. Early literature focused heavily on "time-series studies that attempted to estimate the effects of *NMW*s focusing on the labour force status of teenagers where the employment effects were deemed to be the largest" (Brown, 1999). One common shortcoming of this early literature, as mentioned by Brown et al. (1982), is that "the overwhelming majority of [time-series] studies contained no sensitivity analyses whatsoever." Reviews of the available literature by Brown et al., (1982) and Brown (1999) conclude that estimated reductions in teen employment from a 10% minimum wage increase ranged from 1% to 3%, and the estimates were generally "statistically significant". Interestingly, the lower part of this range was regarded as most plausible as these regressions tended to have the best specification of other explanatory variables, particularly coverage, which varied vastly over state and industry.

Towards the end of the 1970's, there was a shift in emphasis in the literature. Cross-sectional data became more widely available, and studies began to make comparisons on the effect of *NMW* on employment across U.S. state lines. Often this focused on differentials between low-wage and high-wage states, where it was hypothesised that low-wage states

would be more adversely affected by federal minima. As was found with time series data, studies that took a more holistic approach to specifying control variables that affected employment probabilities tended to find less severe effects. Studies from Welch and Cunningham (1978) found a strong adverse effect from minima using the U.S. 1970's census. While Ehrenberg and Marcus (1979) included controls for school expenditure per pupil; adult female education rates and a ratio for the non-white population to the overall population, in their regression found a small positive effect from the same data.

It can be argued that by the mid-1980's, a consensus was forming from time-series and cross-sectional studies that employment effects from *NMW*s were negative but likely small. Nevertheless, his homogenous view has been drawn into question by 1990s studies using longitudinal data. Card and Krueger (1994) examined the employment effects of the New Jersey minimum wage increase by conducting their own interviews of 410 fast food restaurants in New Jersey and Pennsylvania. Employing a difference-in-difference methodology, they observed positive employment effects and found stores in New Jersey increased employment by 13% compared to similar stores in Pennsylvania. Neumark and Washer (1995) disagreed with this conclusion. They directly consulted payroll data, which was deemed to be more accurate, from 230 of the stores included in the Card and Krueger study and estimated the effect of the NJ minimum wage was negative. They found an elasticity of employment with respect to the *NMW* was -0.24 in NJ, while Card and Kruger estimated an elasticity of +0.93.

In the case of legislative minima in the U.K., early studies used the dramatic decline in the toughness of regulation imposed by industry specific Wages Councils throughout the 1980's to determine its effect on employment. Using a constant elasticity of substitution production function, Kaufman (1989) estimated the partial elasticities of substitution in industries subject to statutory minimum wages. Kaufman concluded that "increases in the *NMW*

reduces the employment of women, however, in cases where the price elasticity of demand for output is sufficiently small, male employment is usually unaffected and could even increase." However, using the same New Earning Survey data Dickens et al. (1999) and Machin and Manning (1994) found no evidence of employment reductions. Dickens et al., (1999) postulated that Kaufman's results were due to the fact that his studies concentrate too strongly on "small manufacturing industries and excludes several of the large service-sector industries," notably retail and catering.

Since the U.K. *NMW* took effect, there have been a variety of studies to estimate its effects on low paid employment. Initially studies were conducted using data collected directly by the authors, focusing primarily on low wage sectors particularly vulnerable to minima, such as cleaning and security (Bullock et al., 2001) or residential care homes (Machin et al., 2002) (Machin et al., 2004). These studies report evidence of employment reductions, but are often dismissed as being minor in magnitude and lacking in statistical significance. Criticisms of these studies also arise from the nature of the product market of industries analysed. Machin et al., (2002) argues that "the sector examined is special in that homes are constrained in their ability to pass higher wage costs on into higher prices," and as a result the employment effect estimated does not accurately reflect how the *NMW* impacts the U.K. as a whole. Other studies have drawn frequently from the Labour Force Survey (LFS) to conduct before and after analyses of the employment effects. Stewart (2004) utilizes a difference-in-difference estimation strategy in order to examine the 1999 *NMW* introduction, but finds no probable effect on employment that is statistically significant from zero. One drawback of Stewart's study is the measurement error in the LFS wage variable used to determine those affected by the *NMW*.

Dolton et al. (2012) uses an incremental differences-in-differences methodology to estimate medium and long run impacts of the *NMW* up-ratings over time, focusing closely on the differential impact across heterogeneous geographical

areas. They conclude that an increased bite of the *NMW* is associated with a neutral average employment effect over the entire period, but significant positive *NMW* effects from 2003 onwards, while also finding that the employment rate appears to have risen more in areas where the *NMW* has more relevance. Most recently, Dickens et al. (2012) looked to re-examine the impact on employment on the most vulnerable workers, namely part-time females. They find reductions in employment retention among part-time females, which is further exacerbated by the 2008 recession. Literature on legislative minima in the U.K. has been characterised by debate on whether there is an effect on employment and, if so, what the nature of that effect is. As of now, no in-depth employment analysis of the *NLW* has been conducted. In order to further the literature in this field, it is important to analyse the effect of the *NLW* on employment probabilities of low paid individuals.

3 Estimation Strategy

This methodology aims to estimate the introduction of the minimum wage on the employment probabilities of those affected. The sharp increase in wages caused by the *NLW* and access to individual level longitudinal data makes the difference-in-difference approach an intuitively appealing estimation strategy.

When determining the effect of the *NLW*, the optimal test would be to observe the counterfactual, essentially find what the employment rate of those affected by the *NLW* would have been if the new legislation had not been introduced, ceteris paribus. This information would allow accurate determination of the effect of the *NLW* on employment. Still, the counterfactual and treatment cannot be observed in the same period. The difference-in-difference method attempts to overcome this issue by observing the change in employment of a control group, who act as a proxy for the counterfactual. Therefore, the difference-in-difference approach aims to overcome the issue of what the employment status, of those individuals earning less than the *NLW*, would have been if

the *NLW* had not been introduced.

To give a more precise explanation of the estimation strategy, define eoit to be the employment status of an individual i at time t who is not subject to *NLW* legislation, where eoit=1 if they are employed and eoit =0 if they are not employed. Similarly, let e1it be the employment status of someone who is subject to the *NLW*. It should be noted that only one of these states can be observed by a given individual in a given period. Consider now that the *NLW* is introduced at time t*NLW* , prior to which there is no *NLW*. Classifying individuals into groups g,then for a given group there is direct information on the employment rate for those both affected and unaffected by the *NLW* legislation:

(1) $E(e_{0it} \mid g\ ,\ t)$ for $t < t_{NLW}$ is the employment rate in the absence of the NLW.

(2) $E(e_{1it} \mid g\ ,\ t)$ for $t \geq t_{NLW}$ is the employment rate in the presence of the NLW.

This model is attempting to observe the counterfactual, E(eoit, g , t , t≥t*NLW*), the employment rate in the absence of the *NLW*. This is achieved by making comparisons across groups. Assume that:

(3) $E(e_{0it} \mid g\ ,\ t) = \alpha_g + \gamma_t$

Where g is fixed over time and t is common across groups. Hence it can be assumed that in the absence of the *NLW*, the difference in employment rates between groups is constant in each time period - this is also known as the parallel trends assumption. This key assumption allows for simple difference-in-difference estimates. Assuming that the *NLW* only has an effect on employment probabilities of group 1, then it can be shown that:

(4) $E(e_{1it} \mid g = 1\ ,\ t) = E(e_{0it} \mid g = 1\ ,\ t) + \theta$

(5) $E(e_{1it} \mid g = 2\ ,\ t) = E(e_{0it} \mid g = 2\ ,\ t)$

Considering 2 time periods, where t1 is prior to the *NLW* and t2 is after the implementation of the *NLW*, such that

$t_1 < t_{NLW} t_2$. Double differencing the sample means between groups and across periods gives the simple, or raw, difference-in-difference estimate:

$$(6) \quad [E(e_{it}|\, g = 1\,,\, t_2) - E(e_{it}|\, g = 2\,,\, t_2)] - [E(e_{it}|\, g = 1\,,\, t_1) - E(e_{it}|\, g = 2\,,\, t_1)] = \theta$$

Alternatively, (6) can be rewritten in order to give the employment status of an individual i in a group g at time period t, given the same assumptions as above.

$$(7) \quad e_{it} = \alpha_g + \gamma_t + \theta D_{it} + \varepsilon_{it}$$

Dit represents a dummy variable which is equal to 1 if individual i is affected by the minimum wage (g=1 and $t \geq t_{NLW}$), and is equal to 0 in all other cases and where E(it|g, t)=0. Hence, this regression can be used with individual level longitudinal data to determine the raw difference-in-difference estimator. As discussed in the review of the literature, it is important to control for other variables that affect employment, as the individuals in the control and treatment group may differ in characteristics that mean they are more or less likely to be employed. Failure to control for these differences will lead to omitted variable bias and will adversely affect the validity of the difference-in-difference estimates. This problem can be overcome by extending the difference-in-difference specification and adding a vector of individual characteristics, xit, that are thought to affect the probability of employment. These control variables give the regression adjusted difference-in-difference estimator:

$$(8) \quad e_{it} = x_{it}\beta + \alpha_g + \gamma_t + \theta D_{it} + \varepsilon_{it}$$

Control variables are included to account for differences in characteristics between the treatment and control groups that are not encompassed by the additive and group time effects. This model examines the transitional probability of employment, that is the probability of those currently employed still being employed in the subsequent period as a function of an individual's wage group prior to the increase in the minimum wage. As a result, the economic specification is adjusted from a linear model to a logit model of the form:

$$(9) \quad \Pr Pr\,(e_{it+1} = 1 \mid e_{it} = 1) = \Lambda\{x'_{it}\beta + \alpha_1 g_{1it} + \alpha_3 g_{3it} + \gamma_0 d_{t+1} + \theta g_{1it} d_{t+1} + \varphi g_{3it} d\,(t+1) + \gamma_t\}$$

Notation	Definition
$(e_{it+1} = 1 \mid e_{it} = 1)$	Probability of being employed in a subsequent period given that an individual is employed in the current period.
Λ	The logit transformation. The cumulative distribution function of the logistic distribution.
x_{it}	A vector of factors, other than the National Living Wage, that affects the probability of remaining in employment.
g_{1it}	Binary variable: $= 1$ if: $w_{it} < m_i$ Where w_{it} is the real wage of an individual at time t and m_i is the value of the minimum appropriate to individual i. $= 0$ where else.
g_{3it}	Binary variable: $= 1$ if: $w_{it} > m_i(1 + c)$ $= 0$ where else.
d_{t+1}	Binary variable: $= 1$ if the minimum wage was in place at t+1 $= 0$ where else.
γ_t	Time effects for the remaining time periods.

Figure 3 - Notation of the logit model

Figure 3 details the notation of the logit model. This model will address whether individuals whose wages have had to rise to comply with the *NLW* legislation have a higher or lower probability of being employed in the second period, relative to the comparison group whose wages are just above the new minimum. Note that is still the parameter of interest and the difference-in-difference estimator.

3 Explanation of the Data

3.1 Longitudnal Data

Individual longitudinal data is required to use the difference-in-difference method and must exhibit a number of key characteristics. Firstly, it must include a cross-sectional element that allows a number of characteristics to be observed. Secondly, it should have a time series element that allows these characteristics for individuals to be observed over time. Employment status in the respective 'before' and 'after' *NLW* periods and a robust measure of an individual's

pay rate in the period prior are essential for disaggregation of the control and treatment group, so datasets should also include this information along with other factors that affect the probability of being employed. Data must be constructed such that the observations in periods 1 and 2 straddle the implementation of the *NLW* (1st April 2016). Finally, the datasets must include a large enough number of observations in order to provide robust estimations.

3.2 Labour Force Data

The longitudinal Labour Force Survey (LFS) is a suitable data source. It is collected quarterly, observing individuals for five quarters of data. To match the specification explained above, data from Quarter 2, 2015 to Quarter 1, 2017 can be used. Data will be compiled from 4 different LFSs using observations from before and after the *NLW* in the relevant studies (Figure 4).[1]

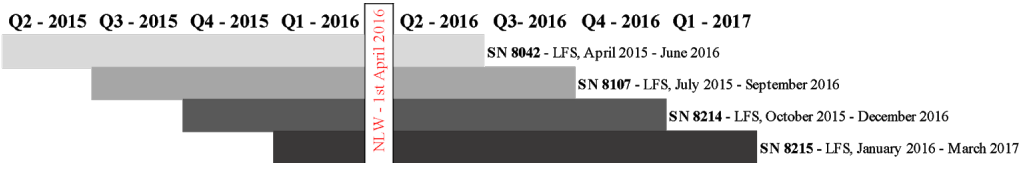


Figure 4 - Dataset Construction Methodology

One limitation of the longitudinal LFS is that hourly pay measurements are only collected in the first quarter of every respective study. For example, SN8042 wage data is collected in Q2 2015 which may affect the reliability of the data due to the fact that wages may have already risen by Q1 2016. Nominal wages have been converted to real wages[2] using the Office of National Statistics Retail Price Index, in order to mitigate the inflationary effect on wages over time. However, the LFS hourly wage data is still subject to measurement error as the survey does not directly consult individuals' payslips for wages but instead constructs this variable with other information. This may lead to measurement issues aforementioned

2 Adjusted for April 2016 prices.

in the debate between Card & Krueger (1994) and Neumark & Washer (1995).

5 Results of Basic Specification

The sample data has been restricted to only include individuals aged between 25 and 65 (65 being the average age of retirement) who were employed in the period of analysis prior to the *NLW*. Those who were either full time students or failed to provide an answer to the interviewer have been omitted. As stated in the discussion of the methodology, the dependent variable in this model is the employment status of the individual after the *NLW* (t+1), given that person was employed in the period prior. Those employed are defined as employees, self-employed, unpaid family workers, and individuals on government employment programmes. If the individuals are not employed in (t+1) they are either unemployed as defined by the Internal Labour Organisation (ILO) or are economically inactive. Initially, the wage variable used to determine those affected by the *NLW* is the LFS variable *HOURPAY* (column A). The treatment group will be defined as those earning less than the *NLW* (£7.20) prior to its implementation and subject to increased wages in order to comply with legislation in the subsequent periods ($65 \geq ageit+1 \geq 25$). The control group will be individuals in the same age demographic but are positioned slightly higher in the wage distribution. Jardim et al. (2017) employ an upper bound of twice the new minimum in their analysis of the effects of legislative minima in Seattle, therefore, the control group will be those earning at least the *NLW* (£7.20) but under £15 (just over double the *NLW*).[3] Appendix 2 gives the raw employment probabilities of individuals in the treatment and control group in the 6 quarters straddling the implementation of the *NLW*. Results presented in Appendix 2 reveal that in periods after the *NLW* the probability of being employed in both the control and treatment groups are lower when compared with the periods prior. It can also be inferred that both groups have a high probability of employment with all 6 periods of analysis reporting over 90% of individuals in work. Raw probabilities

will be useful in determining the magnitude of the effect of the *NLW* on employment.

Table 1 shows the raw linear difference-in-difference coefficients for the effect of the *NLW* over different time periods. Tests [1] and [2] analyse the two quarters that directly straddle the implementation of the *NLW*, Q1 2016 to Q2, 2016. Test [1] includes all individuals that earn less than the *NLW* in the treatment group, and finds a small and statistically significant negative effect of the *NLW* on employment. In Test [2] individuals earning less than, £6.70 (the 2015 *NMW*) in Q1, 2016 are excluded from the treatment group in order to limit the effect of false reporting on the regression estimates. Omitting observations in this range leads to raw coefficients remaining negative but with less statistical significance. Tests [3] and [4] examine extended periods of time in order to determine whether the employment effect of the *NLW* is transitory or persistent. Both tests find slightly larger negative effects of the *NLW* with these effects increasing over time. Furthermore, both results are statistically significant at the 1% level.

The full model is then estimated with control variables. Educational controls are included for whether or not an individual completed: a degree (*GNQ*[4] level 6); another form of higher education below degree level (*GNQ* level 4 or 5) and *GCSE*s or O-Levels (*GNQ* level 2). An individual's experience in their job, determined by the number of years they have been employed with their current employer, and a quadratic term for experience is also encompassed in the control vector. Dummy variables for whether individuals have received job specific training in the last 3 months, their marital status; their region of residency; their sex and whether they have a health related issue or a disability lasting more than 12 months, which will affect the type of work they can do are also contained as controls. Finally, a discrete variable detailing the number of dependants, under the age of 16 the individual has, is included[5]. In each case the controls in the full linear and logit models

4 Government National Qualification.

5 Definitions of LFS variables used in the full model (Appendix 4).

reduce the estimated negative effect of the *NLW* from the raw estimates and are statistically significant to the 5% level in all cases, with the exception of Test [2].

Finally, the full logit model with controls is estimated. In order to interpret the magnitude of the employment effect, the logit coefficient is converted to the "marginal effect" of the dummy variable of interest. Hence, the values for the logit model should be interpreted as the effect of the introduction of the *NLW* on the probability of subsequent employment. For example, those affected by the implementation of the *NLW* in Test [1] have a (-2.455%) lower probability of being employed in Q2, 2016. As reported in Table 1, the logit model generally predicts that the *NLW* lowers employment probabilities for those affected. Logit estimates, in Table 1 are almost always significant at the 1% level with the exception of Test [2] where it is only significant to the 5% level. The large sample size used in these estimates should be kept in consideration when evaluating the significance of the results. Indeed, one could argue that critical values should rise with large sample sizes and, hence significance at the 5% level may be inappropriate for the aforementioned tests. Tests [3] and [4] logit estimates suggest that the magnitude of negative effects of the *NLW* on employment increase over time, to between (-3.213%) and (-3.737%). This would suggest a much larger impact on the employment probabilities of those affected by the *NLW*. Considering that both the treatment and control groups have high raw probabilities of employment, the estimated effects, although not desirable, are not entirely detrimental to the subsequent employment probabilities of the treatment group. Having said this, the estimated negative effects are much larger than previous studies. Future discussions will focus on the logit estimates, although raw and full linear estimates have also been reported in the Results Tables in the interest of comparison.

5.1 How to Construct the Hourly Pay Variable?

Observations of individuals' wages before the *NLW* are es-

sential in distinguishing between individuals who are subject to changing employment pressures (treatment group) and those who are not (control group). A robust measure of the individuals' hourly wages is, therefore, crucial when analysing the effects of the new legislative minima. Prior tests have been conducted using the *LFS*'s derived *HOURPAY* variable. This variable is constructed by dividing gross weekly pay in an individual's main job (GRSSWK) by the sum of their total usual hours worked in their main job (BUSHR) and their usual hours of paid overtime (*POTHR*). This variable is subject to measurement error due to the fact that these are self-reported variables. There is also a disparity between the usual hours worked and the actual hours worked by individuals in the reference period. Dickens and Draca (2005) attempted to overcome this issue by using the *LFS HRRATE* variable, which only gives the wage of individuals who earn based on an hourly rate. Although this is arguably a more accurate measure of an individual's wage, it also leads to a number of observations being omitted from the treatment and control groups for not being able to report an hourly rate. Fortunately, the *LFS* also collects data on actual hours worked in an individual's main job in the reference week (*TTACHR*). This variable excludes those individuals that did not work in the reference week even though they had a job, for example, those on holiday, or sick leave and includes any paid or unpaid overtime. Although this may remove some observations from the analyses, it should provide a more robust measure of actual hours worked by individuals with fluctuating employment over the reference period. Hence, an *ACTUALPAY* (column B) variable is generated by dividing gross weekly pay (*GRSSWK*) by total actuals hours worked (*TTACHR*). It should be noted, however, that the *TTACHR* variable is self-reported, and as such, is still subject to measurement error. Nevertheless, this alternative measure may provide a more insightful estimate of low paid individuals' hourly wages, and provides a useful test of sensitivity and comparison.

Table 1, Column B gives the estimates when the *ACTUALPAY* variable is used to determine the treatment and control groups. In all cases except Test [1] the use of the *ACTUALPAY*

variable is deemed to increase the negative effect of the *NLW* on employment for low wage individuals without drastically changing the statistical significance of estimates. Tests [3] and [4] report significantly larger negative effects compared with the *HOURPAY* tests for the same periods (-4.768% and -4.759% respectively), with the negative effect falling over time. It could, therefore, be argued that the results reported in Column A, underestimate the negative impact on employment, if it is assumed that this alternative wage measure is more accurate. The alternative pay variable draws into question the validity of the basic specification. In the interest of comparison, all subsequent tests and checks of robustness report results when both the *HOURPAY* and *ACTUALPAY* variables are used.

5.2 Regional and Sex Differentials

Often it is argued that the employment effects of the *NLW* will not be felt equally across the UK. Nick Bosanquet (Imperial College London) posits that the *NLW* "will have less impact in London and much more in northern conurbations" because these areas exhibit the greatest productivity problems and the highest incidence of low pay. Conversely, Dolton et al., (2012), found that areas where the *NMW* had the most significance tended to observe positive employment changes. It is, therefore important to determine how the employment effects of the *NLW* differ geographically. In order to determine the regional effects of the *NLW*, regions have been separated into areas of high (20% ≤ %Low Paid), medium (15% < %Low Paid < 20%) and low (%Low Paid ≤ 15%) incidence of low wage employment. This is determined by the percentage of individuals in the region earning less than the *NLW* in quarter 1, 2016 (Appendix 5 and 6). Figures 5 and 6 show that the incidence of low pay varies considerably by region. Dummy variables for these groups are then generated and used to create interaction terms with the variable determining if an individual is affected by the *NLW* in order to ascertain the regional effects.

Table 2 shows the difference-in-difference estimates on the

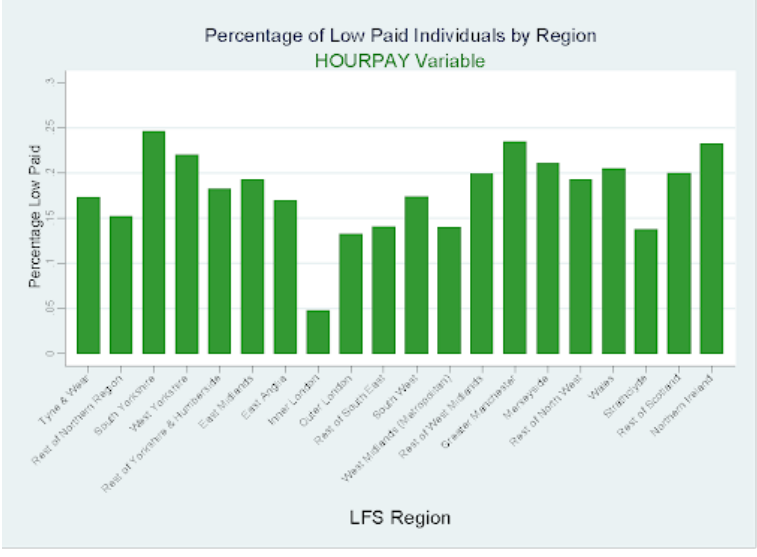


Figure 5 - Percentage of Low Paid Individuals by Region (*HOURPAY* Variable)

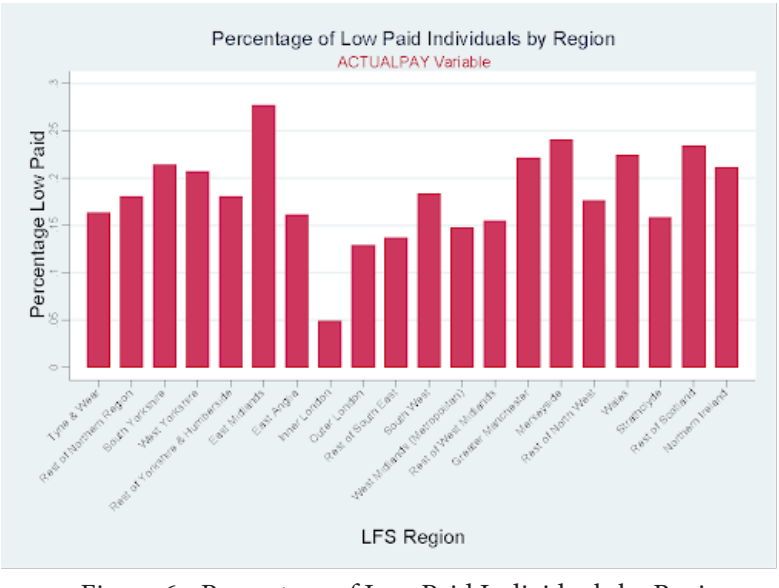


Figure 6 - Percentage of Low Paid Individuals by Region (*ACTUALPAY* Variable)

effect of the introduction of the *NLW* on the probability of subsequent employment, with respect to the incidence of low paid labour in those regions. Tests [5], [6], and [7] examine a range of periods straddling the *NLW*. The most severe negative employment implications appear in regions with a medium incidence of low pay which report a large negative effect of (-6.297%) (Test [7], Column [B]). Medium incidence areas of the country were the only group to find negative results that are statistically significant from zero. Areas of high incidence estimate the effect of *NLW* to be negative, but are statistically insignificant from zero while areas of low incidence saw no statistically significant impact. Interestingly, in Tests [6] and [7], when the *ACTUALPAY* variable is used, areas of low incidence of low pay observe a positive employment impact, despite being statistically insignificant. Therefore, one might concur with the premise that the employment effects of the *NLW* have not been felt evenly across the UK.

Dickens et al. (2012) and Kaufman (1989) found that female employment tended to be affected to a greater extent from legislative minima than male employment. Furthermore, women make up a far greater percentage of those individuals who earn low wages and are subject to wage increases (Figures 7 and 8). It is imperative that the effects for different sexes are determined to see if the new legislation disproportionately affects women. Interaction terms are generated between an individual's sex and the dummy variable of interest, and are included in the regression to determine the effect by sex.

Figure [8] illustrates the estimated effects for men and women across three different periods. Despite *ACTUALPAY* data from Test [8] showing evidence that women are more adversely affected, the rest of the tests draw quite the opposite conclusion. Tests [9] and [10] suggests that men are more negatively affected by the *NLW* legislation. It should be noted that there are fewer observations of men in the sample data. As a result, estimates of the effects on men tend to be statistically insignificant at an appropriate level. Consequently, one should remain cautious in drawing conclusions from these estimates about the impact of the introduction of

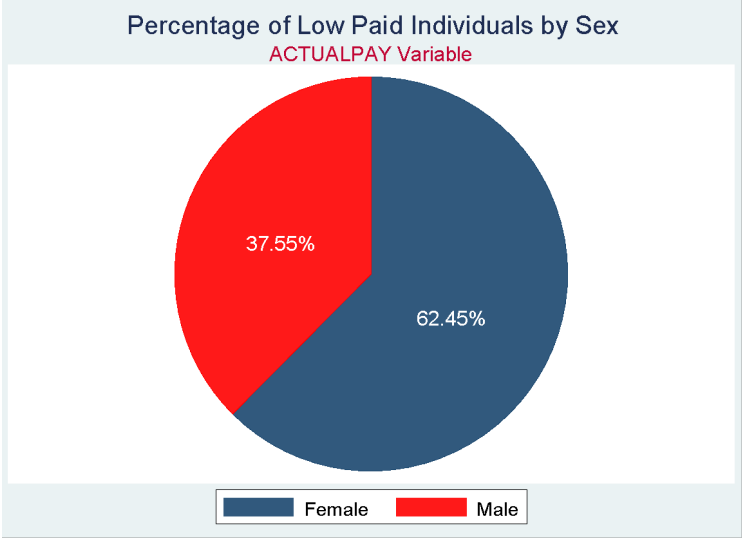


Figure 7 - Percentage of Low Paid Individuals by Sex (*HOURPAY* Variable)

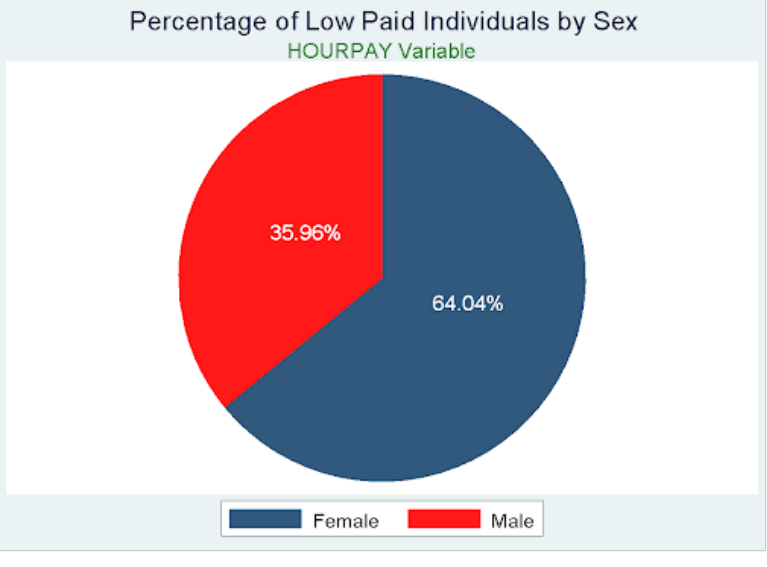


Figure 8 - Percentage of Low Paid Individuals by Region (*ACTUALPAY* Variable)

the *NLW* on differing employment probabilities of men and women. Nonetheless, these tests posit that males incur more adverse employment changes from the *NLW* when compared to females.

6 Tests of Robustness

In order to increase the credibility of the results provided by the basic tests, it is vital that the underlying assumptions of the difference-in-difference methodology are tested.

6.1 Definition of the Control Group

Difference-in-difference estimates rely heavily on the premise that there are significant differences in the treatment and control groups, specifically that the comparison group is not subject to increasing wage pressures as they already earn in excess of *NLW*. One plausible threat to this assumption is caused by the measurement error in the LFS wage variable, referred to in Section 5. The inability to correctly determine an individual's wage may lead to misspecification of those in the control or treatment groups, diminishing the distinction between the two sets of individuals. The dissimilarity between control and treatment groups is also threatened by spill-over effects. For example, the possibility of a cascading wage effect, where wages in the control group rise to maintain the disparity in pay between workers, or increased job search intensity for those earning below the *NLW*. For example, the possibility of a cascading wage effect (i.e. wages in the control group rising in order to maintain the disparity in pay between workers); or increased job search intensity for those who were earning below the *NLW*. In this instance, the actual employment effects may be larger than initially anticipated. Despite Dickens and Manning (2001) finding that "the *NMW* has had virtually no impact on the pay of workers not directly affected," it is important to test the sensitivity of the results with respect to the definition of the control group. As a consequence, this investigation will take the form of altering the bounds of the control group in order to observe if the effect on employment changes.

Changing the definition of the control group can have both advantages and disadvantages to the robustness of the estimates (see Table 4). In Test [11], the upper bound of the wage variable has been increased to £17. By increasing the upper bound for the control group and the robustness of the estimate by reducing the effect of misspecification, both wage spill-overs and substitution between the groups become less likely. Benefits will also be derived, as widening the control group increases the number of observations and in turn the precision of estimation, ceteris paribus. Conversely, increasing the range of the wage distribution captured will compromise the comparability between the groups, as individuals higher in the wage distribution will provide a less comparable counterfactual for the treatment group. Test [12] reduced the upper bound of the wage variable to £13 in order to increase comparability between the groups; but, in the opposite manner to Test [11], it will suffer more from identification errors. Tests [13] and [14] introduce a gap between the minimum wage and the lower limit of the comparison group. Omitting observations in this range should limit misspecification error, but also arguably removes the most comparable individuals from the control group.

Table 4 observes the difference in employment probabilities between quarters 1 and 3, 2016. In all cases when the *ACTUALPAY* wage variable is used, changing the definition of the control group reduces the employment effect of the *NLW* from (-4.768%) to between (-3.791%) and (-4.316%). It should be noted that tests including individuals higher in the wage distribution (Test [11] and [13]) tended to report smaller negative effects than the tests that did not. Most HOURPAY tests also found smaller effects than the basic specification (-3.213%). Test [14] is the clear exception to this, reporting a small increase in the negative effect (-3.271%). Altering the bounds of the control group had no sizeable impact on logit estimates' statistical significance in any of the tests in Table 4. These robustness checks confirm that the effect of the *NLW* is negative on employment, but cast doubts on the magnitude of the effects found in the basic test.

Unlike previous U.K. minimum wage changes, the *NLW* is a premium only applied to individuals aged 25 and over. Meaning 21 to 24-year-olds who were previously subject to the highest *NMW* are not entitled to the £7.20 hourly rate as of April 1st 2016. The decoupling of the 21 to 24-year-olds from the highest *NMW* bracket provides a unique opportunity to observe another control group. The alternative control group can be defined as those who are earning less than the *NLW* prior to its implementation but are too young to receive the mandatory pay increase (by those aged 21 to24). These individuals should share a number of characteristics with the treatment group because they are located in the same position in the wage distribution, but instead differ by age. This comparison group received no increase in the *NMW* until October 2016 and should provide a useful counterfactual. This again provides another validity check and as Meyer (1995) posits, when examining the robustness of the estimates "the more comparison groups the better".

Table 5 gives the estimates when the alternative comparison group is used. Estimates in column [A] are highly comparable in size to those of the initial control group, but do lack the desired statistical significance in all cases except Test [15]. As with Table 1 linear raw estimates vary between (-1.7) and (-3.5) and increase with the duration of the analysis period. However, unlike Table 1 when control variables are included for the full linear and logit models, the difference-in-difference coefficients increase. The logit estimates closely resemble Table 1, with the exception of *ACTUALPAY* Tests [16] and [17], which report effects of a far smaller magnitude. Estimates using the alternative control group lack statistical significance due to small sample size, as only a limited number of observations exist for this control group in the LFS data.

6.2 Distinction of the 'Before' and 'After' Comparision

One of the key features of the difference-in-difference methodology is that two distinct periods straddle the introduction of the *NLW*, that is there is a prior-*NLW* period and a post-*NLW* period, with a clear distinction between them. For

Figure 9 - Percentage of Individuals below the National Living Wage over Time (HOURPAY Variable)

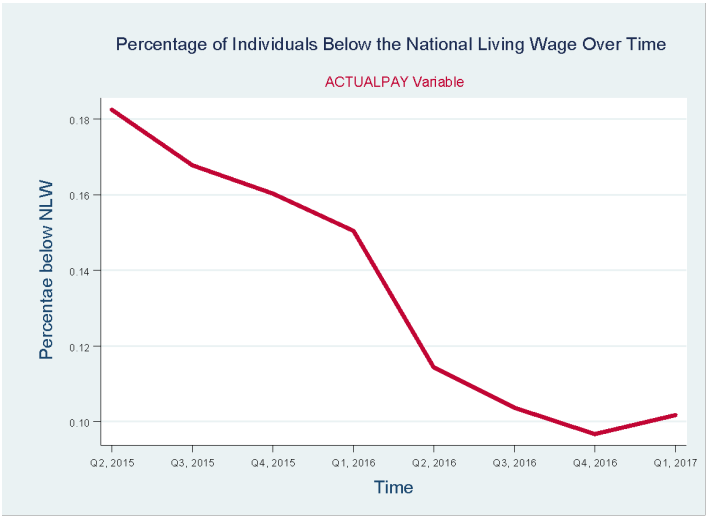


Figure 10 - Percentage of Individuals below the National Living Wage over Time (*ACTUALPAY* Variable)

example, employment status data from Q1, 2016 is treated as unaffected by the new legislative minima, while the subsequent quarters are deemed to be affected. Despite this, it is reasonable to assume that at least some of the changes required to comply with the new minima would be implemented prior to the legally required date, as employers look to change payment practices in order not to breach the new law. If there is a significant level of anticipation by employers, then the 'before' and 'after' distinction will be weakened and estimates of the employment effects of the *NLW* may be incorrect. One approach to determining the level of anticipation is to examine the distribution of wages in the run-up to the *NLW*. This will provide some insight into whether or not wages see a sharp increase prior to the minimum wage, which would provide evidence that the distinction between the two periods is less robust.

Figures 9 and 10 provide the percentage of employees earning below the *NLW* in quarters before and after the *NLW*. Both graphs show the percentage of those who are earning below the national minima is falling in the run-up to the *NLW* and hence diminishes the distinction between the two periods. However, it should be noted that the largest reduction in individuals below the *NLW* in both graphs is between quarter 1 and 2, 2016, which suggests that the *NLW* introduction did have a sizeable effect on low wages. Ergo, these graphs suggest that there is a degree of anticipation in the quarters prior to the imposition of the minimum wage. Figures 9 and 10 also suggest some key difficulties with the robustness of estimates. In both graphs around 10% of workers earned less that £7.20, even as late as quarter 1, 2017. This evidence could again highlight the existence of substantial measurement error in the wage variable, a degree of non-compliance with the *NMW*, or (most likely) a combination of the two. These observations violate the assumptions of the difference-in-difference methodology and weaken the robustness of estimates.

In order to overcome the issue of anticipation, periods prior to the *NLW* can be excluded in order to determine a more robust measure of the policy's effect. Stewart (2004) argues this

'neutral zone' creates a better distinction between the pre- and post-*NLW* periods. Hence various tests are conducted to determine the effect of the policy change when quarter 1, 2016 and/or quarter 4, 2015 are excluded (Table 6). In almost all cases the marginal effects of the logit model are smaller in magnitude but with negative effects, fluctuating in the range of (-2.8%) to (-1.7%). As was found in the basic specification test that examined extended periods, tests [19] and [21] found larger adverse effects on employment probabilities. The creation of a neutral zone had no considerable impact on the statistical significance of the logit estimates. These estimates are surprising because rising wages in Figures 9 and 10 predicted a degree of anticipation by employers that has not translated into an increased negative employment effect in the regression estimates.

6.3 Placebo Test

Another pivotal assumption of the difference-in-difference estimation technique is the parallel trend assumption. It requires that in the absence of treatment (which in this case in the *NLW*), the difference between the treatment and control group is constant over time. This means that the evolution of employment trends for the treatment and control group must be consistent. Alternatively, without the *NLW* reform the trend in employment would have been the same for both groups. If the treatment group has lower probabilities of being employed, then the difference-in-difference estimator may overemphasize the negative effect of the *NLW*, and vice versa. One common robustness or "falsification" test of the common trends assumption is a placebo test. This involves re-estimating the difference-in-differences model over a pre-*NLW* period, but with the assumption that the treatment took effect at an earlier date. Data from 4 LFS longitudinal surveys, straddling the 1st April 2015 has been used in order to conduct the placebo test. This date is selected because it is exactly one year prior to the implementation of the *NLW*. Estimates for this data are taking place prior to the implementation of the *NLW*, thus it is expected that the difference-in-difference estimator should be equal, or at least close to zero,

and statistically insignificant.

Table 7 shows the difference–in-difference estimates over varying controls and time periods that straddle the 1st April 2015. Column [A], Tests [22] and [23] are the only evidence that there is a violation to the common trends assumption finding small (-0.988% and -0.891%, respectively), but statistically significant marginal effects to the coefficients of the logit model. In all other cases, the placebo tests show difference-in-difference estimates that are close to zero with little statistical significance to suggest that they are different from zero. In this case, there is evidence to suggest that the parallel trends assumption for this control and treatment group's holds. This increases the validity of the estimates of the basic specification.

6.4 Did Wages Increase?

Estimating the impact of legislative minima using difference-in-difference relies on the premise that the individuals affected by the *NLW* receive a bigger wage boost than those above the minimum, who represent the control group. If there is a high incidence of employers refusing to raise wages in line with the new legislation or there are cascading wage effects resulting in the control group receiving wage increases, this assumption may be violated. In this case, the difference-in-difference estimates will be biased. Examining the average wages of the treatment and control groups prior to and after the *NLW* provides evidence in support of the aforementioned assumption. Figure 11 shows that the percentage changes in wages for the treatment group are far greater in both nominal and percentage terms when compared with the control group, in the period straddling the minimum wage. One would therefore assume that wage pressures in the treatment group were stronger than the control group and support the validity of estimates.

6.5 Other Robustness Tests

The results prior to this have used a very expansive definition

of those who are not employed as it has included those who are economically inactive. This is the group who is not actively seeking employment and would not accept a job position if

	Average HOURPAY1	Average HOURPAY5	Nominal Change	Percentage Change
Treatment	£6.97	£8.00	£1.03	14.81%
Control	£10.45	£10.83	£0.38	3.62%

	Average ACTUALPAY1	Average ACTUALPAY5	Nominal Change	Percentage Change
Treatment	£6.97	£8.34	£1.37	19.66%
Control	£10.64	£11.16	£0.52	4.89%

Figure 11 - Average Wage Data Prior to and After the National Living Wage

it was offered to them. The three main reasons for economic inactivity are retirement, long term sickness and disabilities, and staying at home to assist a family member in some capacity. Employment changes of this sort are not directly due to the effect of the *NLW* policy, so Table 8 shows the estimates when those defined as inactive in the LFS are removed from the analysis. In both the HOURPAY and *ACTUALPAY* tests spanning different ranges of time, the removal of this group leads to significantly lower estimates of the effect of the *NLW* and also reduces the statistical significance of those estimates. These results suggest the estimates in the basic tests may overestimate the negative employment effects on individuals seeking work.

The basic specification includes individuals aged 25 to 65. The upper bound of this range was selected as it represents the average retirement age of those in the U.K. as measured by the Department of Work and Pensions. However, a large group of individuals will choose to retire prior to this age, and as such will have different employment pressures regardless of the *NLW*. In order to understand the effect the *NLW* will have on younger people, Tests [28] and [29], restrict the age variable to those aged 50 and under in quarter 2, 2016. In

both cases, the magnitude of the logit marginal effects are reduced when this older group is omitted and the statistical significance falls. The fall in significance could be due to the reduced number of observations. The smaller effect suggests that the effect of younger people is smaller than originally thought.

7 Conclusion

This paper utilises LFS longitudinal data to estimate the impacts of the introduction of the *NLW* on the subsequent employment probabilities of those with low wages. This is achieved by adopting a regression adjusted difference-in-difference methodology, which utilises individuals from higher up in the wage distribution or who are too young to be affected by the new legislation as comparison groups. The estimated impact of the *NLW* on the probability of subsequent employment is deemed to be negative, and increases when analysing periods that extend further than the two quarters that directly straddle the implementation of the new minimum. Perhaps counterintuitively, regional estimates show that areas of both high and low incidence of low pay show small, and statistically insignificant from zero, effects on employment, whereas areas of medium incidence report statistically significant negative effects. Contrary to claims from Dickens et al. (2012) and Kaufman (1989), male employment probabilities appear to be affected to a greater extent by the *NLW* than female probabilities. Despite the lack of statistical significance, one should remain cautious when drawing conclusions about whether the sexes are affected differently by the new legislation.

The evidence in this paper stands in direct contradiction to some estimates of the effect of previous U.K. *NMW* increases. Namely, the 1999 introduction (Stewart, 2004) and the 2003 increase (Dickens and Draca, 2005), which show no statistically significant adverse employment effects. This paper more closely resembles estimates from (Machin et al., 2002), (Machin and Wilson, 2004), and most recently (Dickens et

al., 2015), which all obtain modest but negative employment effects from U.K. minima.

The *NLW* represents the largest nominal increase in the U.K. legislative minima since the inception of the *NMW* in 1999. Given the magnitude of the change, further research must be conducted to determine the effect on employment. Research will better advise policy makers in the future on the unintended employment implications that *NMW* increases can have. One clear shortcoming of this paper is that the LFS data used is subject to measurement error in the wage variable. In order for future research to improve the validity of its estimates, researchers must use more robust measures of individuals' wages. Directly consulting individuals' payslips or gaining access to HMRC tax receipts would potentially provide avenues to overcome this issue and should be explored in more depth by future papers.

8 Literature References

Abowd, J.M., Kramarz, F., Margolis, D.N. and Philippon, T., 2000. The tail of two countries: minimum wages and employment in France and the United States (No. 203). IZA Discussion paper series.

Addison, J.T., Blackburn, M.L. and Cotti, C.D., 2012. The Effect of Minimum Wages on Labour Market Outcomes: County-Level Estimates from the Restaurant-and-Bar Sector. British Journal of Industrial Relations, 50(3), pp.412-435.

Addison, J.T., Blackburn, M.L. and Cotti, C., 2008. New estimates of the effects of minimum wages in the US retail trade sector.

Brown, C., 1999. Minimum wages, employment, and the distribution of income. Handbook of Labor Economics, 3, pp.2101-2163.

Brown, C., Gilroy, C. and Kohen, A., 1982. The effect of the

minimum wage on employment and unemployment. Journal of Economic Literature, 20(2), pp.487-528.

Bullock, A., Hughes, A. and Wilkinson, F., 2001. The impact of the National Minimum Wage on small and medium sized businesses in the cleaning and security sectors. ESRC Centre for Business Research.

Card, D. and A. B. Krueger 1994. Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. The American Economic Review 84(4): 772- 793.

Card, D., 1992. Using regional variation in wages to measure the effects of the federal minimum wage. ILR Review, 46(1), pp.22-37.

Department for Business, Innovation and Skills. (2015). National Minimum Wage - Low Pay Commission Remit 2016. Available: https://www.gov.uk/government/publications/national-minimum-wage-low-pay-commission-remit-2016. Last accessed 10th November 2017.

Dickens, R., Riley, R. and Wilkinson, D., 2015. A Re-examination of the Impact of the UK National Minimum Wage on Employment. Economica, 82(328), pp.841-864.

Dickens, R. and Draca, M., 2005. The employment effects of the October 2003 increase in the national minimum wage (No. CEPDP0). Centre for Economic Performance (CEP), The London School of Economics and Political Science (LSE).

Dickens, Richard and Alan Manning (2001), "The national minimum wage and wage inequality", mimeo, Centre for Economic Performance, LSE, November.

Dickens, R., Machin, S. and Manning, A., 1999. The effects of minimum wages on employment: Theory and evidence from Britain. Journal of Labor Economics, 17(1), pp.1-22.
Dolton, P., Bondibene, C.R. and Wadsworth, J., 2012. Employ-

ment, Inequality and the UK National Minimum Wage over the Medium-Term. Oxford Bulletin of Economics and Statistics, 74(1), pp.78-106.

Ehrenberg, R.G. and Marcus, A., 1979. Minimum Wage Legislation and the Educational Decisions of Youth. forthcoming in Research in Labor mics.

Freeman, Richard (1982), "Economic determinants of geographic and individual variation in the labor market position of young persons", in: R. Freeman and D. Wise, eds., The youth labor market problem: its nature, causes and consequences (NBER and University of Chicago Press, Chicago, 1L) pp. 115-148.

Georgiadis, A.P., 2006. Is the Minimum Wage Efficient?: Evidence of the Effects of the UK National Minimum Wage in the Residential Care Homes Sector. Centre for Market and Public Organisation, University of Bristol.

Jardim, E., Long, M.C., Plotnick, R., van Inwegen, E., Vigdor, J. and Wething, H., 2017. Minimum Wage Increases, Wages, and Low-Wage Employment: Evidence from Seattle (No. w23532). National Bureau of Economic Research.

Kaufman, R.T., 1989. The effects of statutory minimum rates of pay on employment in Great Britain. The Economic Journal, 99(398), pp.1040-1053.

Kosters, M. and Welch, F., 1972. The effects of minimum wages on the distribution of changes in aggregate employment. The American Economic Review, 62(3), pp.323-332.

Linneman, P., 1982. The economic impacts of minimum wage laws: a new look at an old question. Journal of Political Economy, 90(3), pp.443-469.

Living Wage Foundation. 2017. NEWS: Women Continue to be Hit Hardest by Low Wages in UK. [ONLINE] Available at: https://www.livingwage.org.uk/news/news-women-contin-

ue-be-hit-hardest-low-wages-uk. [Accessed 13 April 2018].

Machin, S.J. and Wilson, J., 2004. Minimum wages in a low-wage labour market: care homes in the UK.

Machin, S., Manning, A. and Rahman, L., 2002. Where the minimum wage bites hard: the introduction of the UK national minimum wage to a low wage sector. Centre for Economic Performance, London School of Economics and Political Science.

Machin, S. and Manning, A., 1994. The effects of minimum wages on wage dispersion and employment: Evidence from the UK Wages Councils. ILR Review, 47(2), pp.319-329.

Metcalf, D., 2008. Why has the British national minimum wage had little or no impact on employment?. Journal of Industrial Relations, 50(3), pp.489-512.

Meyer, Bruce (1995), "Natural and Quasi-experiments in Economics", Journal of Business and Economic Statistics, 13, 151-161

Neumark, D., Schweitzer, M. and Wascher, W., 2004. Minimum wage effects throughout the wage distribution. Journal of Human Resources, 39(2), pp.425-450.

Neumark, D. and Wascher, W., 1995. The effect of New Jersey's minimum wage increase on fast-food employment: a re-evaluation using payroll records (No. w5224). National Bureau of Economic Research.

Pohlman, J.T. and Leitner, D.W., 2003. A comparison of ordinary least squares and logistic regression.

Ragan, J.F., 1977. Minimum wages and the Youth Labor Market. The Review of Economics and Statistics, pp.129-136.

Sarah O'Connor. (2016). Britain's national living wage — economists debate pros and cons. Available: https://www.

ft.com/content/361764ba-f732-11e5-96db-fc683b5e52db Last accessed 10th November 2017.

Stewart, M.B., 2004. The Impact of the Introduction of the UK Minimum Wage on the Employment Probabilities of Low-Wage Workers. Journal of the European Economic Association, 2(1), pp.67-97.

Welch, F. and Cunningham, J., 1978. Effects of minimum wages on the level and age composition of youth employment. The Review of Economics and Statistics, pp.140-145.

West, E.G. and McKee, M., 1980. Monopsony and" Shock" Arguments for Minimum Wages. Southern Economic Journal, pp.883-891.

9 Data References

Office of National Statistics. Retail Price Inflation, UK: March 2018, Time series: RPI All Items Index: Jan 1987=100, https://www.ons.gov.uk/economy/inflationandpriceindices/timeseries/chaw/mm23

Office for National Statistics. Social Survey Division, Northern Ireland Statistics and Research Agency. Central Survey Unit. (2017). Labour Force Survey Five-Quarter Longitudinal Dataset, January 2016 - March 2017. [data collection]. UK Data Service. SN: 8215, http://doi.org/10.5255/UKDA-SN-8215-1

Office for National Statistics. Social Survey Division. (2017). Labour Force Survey Five-Quarter Longitudinal Dataset, October 2015 - December 2016. [data collection]. UK Data Service. SN: 8214, http://doi.org/10.5255/UKDA-SN-8214-1

Office for National Statistics. Social Survey Division. (2017). Labour Force Survey Five-Quarter Longitudinal Dataset, July 2015 - September 2016. [data collection]. 2nd Edition. UK Data Service. SN: 8107, http://doi.org/10.5255/UKDA-SN-8107-2

Office for National Statistics. Social Survey Division. (2017). Labour Force Survey Five-Quarter Longitudinal Dataset, April 2015 - June 2016. [data collection]. 2nd Edition. UK Data Service. SN: 8042, http://doi.org/10.5255/UKDA-SN-8042-3

Office for National Statistics. Social Survey Division, Northern Ireland Statistics and Research Agency. Central Survey Unit. (2017). Labour Force Survey Five-Quarter Longitudinal Dataset, January 2015 - March 2016. [data collection]. 2nd Edition. UK Data Service. SN: 7988, http://doi.org/10.5255/UKDA-SN-7988-2

Office for National Statistics. Social Survey Division. (2017). Labour Force Survey Five-Quarter Longitudinal Dataset, October 2014 - December 2015. [data collection]. 3rd Edition. UK Data Service. SN: 7905, http://doi.org/10.5255/UKDA-SN-7905-3

Office for National Statistics. Social Survey Division. (2017). Labour Force Survey Five-Quarter Longitudinal Dataset, July 2014 - September 2015. [data collection]. 3rd Edition. UK Data Service. SN: 7845, http://doi.org/10.5255/UKDA-SN-7845-3

Office for National Statistics. Social Survey Division. (2017). Labour Force Survey Five-Quarter Longitudinal Dataset, April 2014 - June 2015. [data collection]. 3rd Edition. UK Data Service. SN: 7790, http://doi.org/10.5255/UKDA-SN-7790-4

10 Results Tables

10.1 Table 1: Basic Tests

		Hourpay Test [A]		Actualpay Test [B]	
[1]	**Q1, 2016 to Q2, 2016 *(No lower bound on wage variable)***				
	Raw Linear	-0.02555	***	-0.02027	***
		0.00591		*0.00586*	
	Full Linear	-0.01977	***	-0.01579	***
		0.00563		*0.00561*	
	Full Logit	-0.02455	***	-0.01962	***
		0.00584		*0.00577*	
[2]	**Q1, 2016 to Q2, 2016 *(£6.70 lower bound on wage variable)***				
	Raw Linear	-0.01909	**	-0.03214	**
		0.00937		*0.01289*	
	Full Linear	-0.01352		-0.02736	**
		0.00910		*0.01248*	
	Full Logit	-0.01779	**	-0.03140	**
		0.00898		*0.01273*	
[3]	**Q1, 2016 to Q3, 2016 *(£6.70 lower bound on wage variable)***				
	Raw Linear	-0.03490	***	-0.04783	***
		0.01200		*0.01364*	
	Full Linear	-0.02499	**	-0.03934	***
		0.01153		*0.01297*	
	Full Logit	-0.03213	***	-0.04768	***
		0.01143		*0.01351*	
[4]	**Q1, 2016 to Q4, 2016 *(£6.70 lower bound on wage variable)***				
	Raw Linear	-0.03944	***	-0.04867	***
		0.01252		*0.01410*	
	Full Linear	-0.02779	**	-0.03779	***
		0.01203		*0.01335*	
	Full Logit	-0.03737	***	-0.04759	***
		0.01208		*0.01372*	

*** Statistically Significant at the 1% Level
** Statistically Significant at the 5% Level
* Statistically Significant at the 10% Level
Robust Standard Errors

10.2 Table 2: Regional Tests

		Hourpay Test [A]			Actualpay Test [B]		
		High	Medium	Low	High	Medium	Low
[5]	**Q1, 2016 to Q2, 2016 *(£6.70 lower bound on wage variable)***						
	Full Linear	-0.00016	-0.03553 *	-0.00345	-0.01355	-0.02463 **	-0.00367
		0.01054	*0.01860*	*0.01594*	*0.00976*	*0.00965*	*0.00822*
	Full Logit	-0.00425	-0.03469 **	-0.00894	-0.01992	-0.02663 ***	-0.01023
		0.01249	*0.01669*	*0.01674*	*0.01222*	*0.00991*	*0.01060*
[6]	**Q1, 2016 to Q3, 2016 *(£6.70 lower bound on wage variable)***						
	Full Linear	-0.00774	-0.05102 **	-0.00909	-0.03023	-0.06627 ***	0.00610
		0.01501	*0.02098*	*0.02066*	*0.02522*	*0.02534*	*0.01061*
	Full Logit	-0.01557	-0.04888 ***	-0.01796	-0.03819	-0.05945 ***	0.00019
		0.01374	*0.01503*	*0.01720*	*0.02757*	*0.02235*	*0.01240*
[7]	**Q1, 2016 to Q4, 2016 *(£6.70 lower bound on wage variable)***						
	Full Linear	-0.01403	-0.05770 ***	0.00016	-0.03003	-0.06986 ***	0.01164
		0.01693	*0.02216*	*0.01704*	*0.02604*	*0.02584*	*0.00900*
	Full Logit	-0.02272	-0.05729 ***	-0.01127	-0.03690	-0.06297 ***	0.00527
		0.01838	*0.02063*	*0.01947*	*0.0282*	*0.02313*	*0.01036*

*** Statistically Significant at the 1% Level
** Statistically Significant at the 5% Level
* Statistically Significant at the 10% Level
Robust Standard Errors

10.3 Table 3: Sex Tests

		Hourpay Test [A]		Actualpay Test [B]	
		Female	Male	Female	Male
[8]	**Q1, 2016 to Q2, 2016 *(£6.70 lower bound on wage variable)***				
	Full Linear	-0.01820	-0.02699	-0.01997 ***	-0.01735 *
		0.01965	*0.01136*	*0.00702*	*0.00976*
	Full Logit	-0.01928	-0.02361	-0.02122 ***	-0.01924 *
		0.01180	*0.01716*	*0.00759*	*0.01077*
[9]	**Q1, 2016 to Q3, 2016 *(£6.70 lower bound on wage variable)***				
	Full Linear	-0.03117 **	-0.04137 *	-0.04167 **	-0.04617 *
		0.01391	*0.02240*	*0.01752*	*0.02627*
	Full Logit	-0.03248 **	-0.03413 *	-0.04039 ***	-0.04186 *
		0.01887	*0.01416*	*0.01706*	*0.02299*
[10]	**Q1, 2016 to Q4, 2016 *(£6.70 lower bound on wage variable)***				
	Full Linear	-0.02986 **	-0.05978 **	-0.03935 **	-0.05673 *
		0.01311	*0.02682*	*0.01712*	*0.02957*
	Full Logit	-0.03207 **	-0.04814 **	-0.03850 **	-0.04858 *
		0.01381	*0.02254*	*0.01667*	*0.02584*

*** Statistically Significant at the 1% Level
** Statistically Significant at the 5% Level
* Statistically Significant at the 10% Level
Robust Standard Errors

10.4 Table 4: Alternative Definitions of the Control Group

	Hourpay Test [A]		Actualpay Test [B]	
[11] Q1, 2016 to Q3, 2016 *(£17 upper wage bound on control group)*				
Raw Linear	-0.03396	***	-0.04309	***
	0.01199		*0.01472*	
Full Linear	-0.02233	**	-0.03424	**
	0.01155		*0.01406*	
Full Logit	-0.02989	***	-0.03856	***
	0.01105		*0.01342*	
[12] Q1, 2016 to Q3, 2016 *(£13 upper wage bound on control group)*				
Raw Linear	-0.03427	***	-0.04348	***
	0.01202		*0.01474*	
Full Linear	-0.02499	**	-0.03717	***
	0.01152		*0.01401*	
Full Logit	-0.03271	***	-0.04316	***
	0.01173		*0.01452*	
[13] Q1, 2016 to Q3, 2016 *(£9.20 - £17 wage range on control group)*				
Raw Linear	-0.03597	***	-0.04321	***
	0.01201		*0.01475*	
Full Linear	-0.02295	**	-0.03196	**
	0.01151		*0.01391*	
Full Logit	-0.02896	***	-0.03791	***
	0.01054		*0.01328*	
[14] Q1, 2016 to Q3, 2016 *(£9.20 - £15 wage range on control group)*				
Raw Linear	-0.03542	***	-0.04344	***
	0.01203		*0.01477*	
Full Linear	-0.02253	*	-0.03295	**
	0.01147		*0.01383*	
Full Logit	-0.03037	***	-0.04046	***
	0.01102		*0.01392*	

*** Statistically Significant at the 1% Level
** Statistically Significant at the 5% Level
* Statistically Significant at the 10% Level
Robust Standard Errors

10.5 Table 5: Alternative Control Group

	Hourpay Test [A]		Actualpay Test [B]	
[15] Q1, 2016 to Q2, 2016 *(No lower bound on wage variable)*				
Raw Linear	-0.02506	***	-0.01778	***
	0.00637		*0.00640*	
Full Linear	-0.02689	***	-0.01918	***
	0.00603		*0.00590*	
Full Logit	-0.02785	***	-0.01959	***
	0.00638		*0.00620*	
[16] Q1, 2016 to Q2, 2016 *(£6.70 lower bound on wage variable)*				
Raw Linear	-0.01490		-0.02280	
	0.01098		*0.01394*	
Full Linear	-0.01620		-0.02466	*
	0.01000		*0.01257*	
Full Logit	-0.01733		-0.02481	*
	0.01073		*0.01341*	
[17] Q1, 2016 to Q3, 2016 *(£6.70 lower bound on wage variable)*				
Raw Linear	-0.02956	**	-0.03430	**
	0.01275		*0.01566*	
Full Linear	-0.03139	***	-0.03612	***
	0.00600		*0.01365*	
Full Logit	-0.03242	**	-0.03604	**
	0.01251		*0.01497*	

*** Statistically Significant at the 1% Level
** Statistically Significant at the 5% Level
* Statistically Significant at the 10% Level
Robust Standard Errors

10.6 Table 6: 'Before' and 'After' Comparison

	Hourpay Test [A]	Actualpay Test [B]
[18] Q4, 2015 to Q2, 2016 *(£6.70 lower bound on wage variable)*		
Raw Linear	-0.03626 **	-0.03994 **
	0.01453	*0.01688*
Full Linear	-0.02502 *	-0.02979 *
	0.01375	*0.01582*
Full Logit	-0.01811 ***	-0.01791 ***
	0.00507	*0.00513*
[19] Q4, 2015 to Q3, 2016 *(£6.70 lower bound on wage variable)*		
Raw Linear	-0.04672 ***	-0.04860 ***
	0.01636	*0.01833*
Full Linear	-0.03105 **	-0.03508 **
	0.01539	*0.01700*
Full Logit	-0.02387 ***	-0.02362 ***
	0.00622	*0.00624*
[20] Q3, 2015 to Q2, 2016 *(£6.70 lower bound on wage variable)*		
Raw Linear	-0.04183 **	-0.04643 **
	0.01906	*0.02203*
Full Linear	-0.02423	-0.02900
	0.01736	*0.02025*
Full Logit	-0.02352 ***	-0.02221 ***
	0.00751	*0.00760*
[21] Q3, 2015 to Q3, 2016 *(£6.70 lower bound on wage variable)*		
Raw Linear	-0.05311 **	-0.05223 **
	0.02273	*0.02357*
Full Linear	-0.03115	-0.03124
	0.01999	*0.02099*
Full Logit	-0.02894 ***	-0.02564 ***
	0.00897	*0.00842*

*** Statistically Significant at the 1% Level
** Statistically Significant at the 5% Level
* Statistically Significant at the 10% Level
Robust Standard Errors

10.7 Table 7: Placebo Test

	Hourpay Test [A]	Actualpay Test [B]
[22] Q1, 2015 to Q2, 2015 *(No lower bound on wage variable)*		
Raw Linear	-0.01461 ***	-0.00463
	0.00417	*0.00292*
Full Linear	-0.00850 **	-0.00331
	0.00396	*0.00288*
Full Logit	-0.00988 ***	-0.00339 **
	0.00224	*0.00157*
[23] Q1, 2015 to Q2, 2015 *(£6.70 lower bound on wage variable)*		
Raw Linear	-0.01572 *	-0.00351
	0.00860	*0.00551*
Full Linear	-0.00984	-0.00177
	0.00832	*0.00553*
Full Logit	-0.00891 ***	-0.00235
	0.00339	*0.00239*
[24] Q1, 2015 to Q3, 2015 *(£6.70 lower bound on wage variable)*		
Raw Linear	-0.01963 **	-0.00012
	0.00968	*0.00324*
Full Linear	-0.00963	0.00230
	0.00914	*0.00339*
Full Logit	-0.00207	-0.00038
	0.10136	*0.00326*
[25] Q1, 2015 to Q4, 2015 *(£6.70 lower bound on wage variable)*		
Raw Linear	-0.01791 *	0.00036
	0.00925	*0.00251*
Full Linear	-0.00715	0.00239
	0.00876	*0.00265*
Full Logit	-0.01641	0.00008
	0.08704	*0.00296*

*** Statistically Significant at the 1% Level
** Statistically Significant at the 5% Level
* Statistically Significant at the 10% Level
Robust Standard Errors

10.8 Table 8: Other Robustness Tests

	Hourpay Test [A]	Actualpay Test [B]
[26] Q1, 2016 to Q3, 2016 (*Economic Inactive Removal Test*)		
Raw Linear	-0.00952	-0.03800
	0.00615	*0.02889*
Full Linear	-0.00695	-0.03581
	0.00615	*0.02860*
Full Logit	-0.00881	-0.02956
	0.00594	*0.02513*
[27] Q1, 2016 to Q4, 2016 (*Economic Inactive Removal Test*)		
Raw Linear	-0.01452 *	-0.01667 *
	0.00769	*0.00871*
Full Linear	-0.01089	-0.01457 *
	0.00770	*0.00860*
Full Logit	-0.01396 *	-0.01653 *
	0.00752	*0.00869*
[28] Q1, 2016 to Q3, 2016 (*Restricting the Age Variable to 50*)		
Raw Linear	-0.02343 **	-0.03627 ***
	0.01155	*0.01383*
Full Linear	-0.01913 *	-0.03288 **
	0.01142	*0.01358*
Full Logit	-0.02241 *	-0.03265 **
	0.01159	*0.01264*
[29] Q1, 2016 to Q4, 2016 (*Restricting the Age Variable to 50*)		
Raw Linear	-0.03389 **	-0.03924 ***
	0.01470	*0.01503*
Full Linear	-0.02877 **	-0.03390 **
	0.01451	*0.01472*
Full Logit	-0.03029 **	-0.03330 **
	0.01422	*0.01352*

*** Statistically Significant at the 1% Level
** Statistically Significant at the 5% Level
* Statistically Significant at the 10% Level
Robust Standard Errors

11 Appendix

11.1 Appendix 1: Summary Statistics

Summary Statistics	Number of Observations
Total Number of Individuals	16387
Male	7718
Female	8669
Under 25	14978
Over 25	1409

Note:
Based on 4 Labour Force Surveys from Quarter 2, 2015, to Quarter 1, 2017.
Number of observations will vary from the basic specification depending on changes to the definition of the control and treatment group.
Age as of Quarter 2, 2016.

11.2 Appendix 2: Raw Employment Probabilities

Raw Employment Probabilitites	Q3, 2015	Q4, 2015	Q1, 2016	Q2, 2016	Q3, 2016	Q4, 2016
Control	0.9524	0.9457	0.9507	0.9260	0.9096	0.9107
Treatment Group	0.9814	0.9737	0.9773	0.9661	0.9668	0.9667
Alternative Treatment Group	0.9607	0.8873	0.8989	0.8651	0.9206	0.8684

Note: Quarter 3, 2015 to Quarter 1, 2016 represent periods prior to the National Living Wage.
Quarter 2, 2016 to Quarter 4, 2016 represent periods after the National Living Wage.

11.3 Appendix 3: Definitions of Treatment and Control Groups

Test	Treatment Group	Control Group	Period of Analysis
Table 5: Basic Tests			
[1]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage* in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016	Q1, 2016 to Q2, 2016
[2]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016	Q1, 2016 to Q2, 2016
[3]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016	Q1, 2016 to Q3, 2016
[4]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016	Q1, 2016 to Q4, 2016
Table 8: Regional Tests			
[5]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£9.20 in Q1, 2016	Q1, 2016 to Q2, 2016
[6]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£9.20 in Q1, 2016	Q1, 2016 to Q3, 2016
[7]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£9.20 in Q1, 2016	Q1, 2016 to Q4, 2016
Table 9: Sex Tests			
[8]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£9.20 in Q1, 2016	Q1, 2016 to Q2, 2016
[9]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£9.20 in Q1, 2016	Q1, 2016 to Q3, 2016
[10]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£9.20 in Q1, 2016	Q1, 2016 to Q4, 2016
Table 10: Alternative Definition of the Control Group			
[11]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £17≥*wage*≥£7.20 in Q1, 2016	Q1, 2016 to Q3, 2016
[12]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £13≥*wage*≥£7.20 in Q1, 2016	Q1, 2016 to Q3, 2016
[13]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £17≥*wage*≥£9.20 in Q1, 2016	Q1, 2016 to Q3, 2016
[14	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£9.20 in Q1, 2016	Q1, 2016 to Q3, 2016
Table 11: Alternative Control Group			
[15]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage* in Q1, 2016	24≥*age*≥21 in Q2, 2016 £7.20≥*wage* in Q1, 2016	Q1, 2016 to Q2, 2016
[16]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	24≥*age*≥21 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	Q1, 2016 to Q2, 2016
[17]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	24≥*age*≥21 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	Q1, 2016 to Q3, 2016
Table 12: 'Before' and 'After' Comparison			
[18]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016	Q4, 2015 to Q2, 2016
[19]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016	Q4, 2015 to Q3, 2016
[20]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016	Q3, 2015 to Q2, 2016
[21]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016	Q3, 2015 to Q3, 2016

Table 13: Placebo Test			
[22]	65≥*age*≥25 in Q2, 2015 £7.20≥*wage* in Q1, 2015	65≥*age*≥25 in Q2, 2015 £15≥*wage*≥£7.20 in Q1, 2015	Q1, 2015 to Q2, 2015
[23]	65≥*age*≥25 in Q2, 2015 £7.20≥*wage*≥£6.70 in Q1, 2015	65≥*age*≥25 in Q2, 2015 £15≥*wage*≥£7.20 in Q1, 2015	Q1, 2015 to Q2, 2015
[24]	65≥*age*≥25 in Q2, 2015 £7.20≥*wage*≥£6.70 in Q1, 2015	65≥*age*≥25 in Q2, 2015 £15≥*wage*≥£7.20 in Q1, 2015	Q1, 2015 to Q3, 2015
[25]	65≥*age*≥25 in Q2, 2015 £7.20≥*wage*≥£6.70 in Q1, 2015	65≥*age*≥25 in Q2, 2015 £15≥*wage*≥£7.20 in Q1, 2015	Q1, 2015 to Q4, 2015

Table 14: Other Robustness Tests			
[26]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016 Economically Inactive Removed	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016 Economically Inactive Removed	Q1, 2016 to Q3, 2016
[27]	65≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016 Economically Inactive Removed	65≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016 Economically Inactive Removed	Q1, 2016 to Q4, 2016
[28]	50≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	50≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016	Q1, 2016 to Q3, 2016
[29]	50≥*age*≥25 in Q2, 2016 £7.20≥*wage*≥£6.70 in Q1, 2016	50≥*age*≥25 in Q2, 2016 £15≥*wage*≥£7.20 in Q1, 2016	Q1, 2016 to Q4, 2016

11.4 Appendix 4: Labour Force Survey Variables

LFS Variable	LFS Definition	Expression in LFS	Expression in constructed Dataset.
PERSID	Unique identifier	11 digit number identifying individuals in the 5 Quarter Labour Force Survey Data.	The LFS tracks individuals for up to 5 quarters and this variable is used to identify information on individuals across time.
AGE*	Age of respondent	Numerical form, for all individuals.	Used to determine which individuals would be subject to the NLW legislation.
ILODEFR*	Basic economic activity	Derived Variable which explains the economic activity status of the individual. Reported as: (1) In employment (2) ILO unemployed (3) Inactive (4) Under 16	Under 21s are not measured in the control or treatment group, therefore (4) is removed from the analysis. Employed (=1), ILO unemployed (=0), Inactive (=0).
HOURPAY*	Average gross hourly pay	Constructed by LFS by dividing the Gross weekly wage (GRSSWK*) by the basic usual hours worked (BUSHR*). Reported as amount in pounds (£).	Used to determine those individuals subject to wage increases from the NLW.
GRSSWK*	Gross weekly pay in main job	Reported as gross income in pounds (£) in reference week.	Used by author in the construction of alternative hourly wage variable ACTUALPAY*.
TTACHR*	Actual hours in main job	Reported as hours of work in a reference week, cannot exceed 97 hours.	Used by author in the construction of alternative hourly wage variable ACTUALPAY*.
SEX	Sex of respondent	Reported by the individual, (=1) if male and (=2) if female in LFS data.	Binary dummy control variable. Male (=0), Female (=1).
QUAL_1*	Whether degree level education obtained	Reported by the individual, Yes (=1) and No (=0).	Binary dummy control variable. Yes (=1), No (=0).
QUAL_9*	Whether other higher education qualification below degree level obtained	Reported by the individual, Yes (=1) and No (=0).	Binary dummy control variable. Yes (=1), No (=0).

QUAL 21*	Whether GCSE/Vocational GCSE obtained	Reported by the individual, Yes (=1) and No (=0).	Binary dummy control variable. Yes (=1), No (=0).
ED13WK1*	Job related training or education in the last 3 months	Reported by the individual, Yes (=1) and No (=2).	Binary dummy control variable. Yes (=1), No (=0).
CONMPY*	Year started working for current employer	Reported as the year e.g. 2000.	Used to construct individuals experience variable as (2016-Value of CONMPY*).
MARSTA*	Marital status	Reported as: (1) Single, never married (2) Married, living with husband/wife (3) Married, separated from husband/wife (4) Divorced (5) Widowed (6) A civil partner in a legally-recognised Civil Partnership (7) In a legally-recognised Civil Partnership and separated from his/her civil partner (8) Formerly a civil partner, the Civil Partnership now legally dissolved (9) A surviving civil partner: his/her partner having since died	Binary dummy control variable. Married, living with husband/wife, A civil partner in a legally-recognised Civil Partnership (=1), all other (=0).
FDPCH16*	Number of dependent children in family aged under 16	(0-10) Number of dependent children in family aged under 16.	Discrete dummy control variable.
LNGLST*	Health problems lasting or expected to last more than 1 year	Reported as: (1) Yes (2) No (3) Don't know (4) Refusal	Binary dummy control variable. Yes (=1), all other (=0).
GORWKR*	Region of place of work in main job	Reported as: (1) Tyne & Wear (2) Rest of North East (3) Greater Manchester (4) Merseyside (5) Rest of North West (6) South Yorkshire (7) West Yorkshire (8) Rest of Yorkshire & Humberside (9) East Midlands (10) West Midlands (met county) (11) Rest of West Midlands (12) East of England (13) Central London (14) Inner London (15) Outer London (16) South East (17) South West (18) Wales (19) Strathclyde (20) Rest of Scotland (21) Northern Ireland (22) Workplace outside UK	Reduced to 3 categories based on the incidence of low paid workers as a percentage of total employees.

11.5 Appendix 5: Definition of Regional Breakdown (*HOURPAY*)

LFS ID	Region	Percentage Low Paid
1	Tyne & Wear	17.3%
2	Rest of Northern Region	15.2%
3	South Yorkshire	24.6%
4	West Yorkshire	22.1%
5	Rest of Yorkshire & Humberside	18.3%
6	East Midlands	19.2%
7	East Anglia	17.0%
8	Inner London	4.8%
9	Outer London	13.3%
10	Rest of South East	14.1%
11	South West	17.4%
12	West Midlands (Metropolitan)	14.0%
13	Rest of West Midlands	19.9%
14	Greater Manchester	23.5%
15	Merseyside	21.1%
16	Rest of North West	19.3%
17	Wales	20.5%
18	Strathclyde	13.8%
19	Rest of Scotland	20.0%
20	Northern Ireland	23.3%

Incidence of Low Pay	Definintion	Regions
High	20 ≤ (% Low Paid)	South Yorkshire West Yorkshire Greater Manchester Merseyside Wales Rest of Scotland Northern Ireland
Medium	15 < (% Low Paid) < 20	Tyne & Wear Rest of Northern Region Rest of Yorkshire & Humberside East Midlands East Anglia South West Rest of West Midlands Rest of North West
Low	(% Low Paid) ≤ 15	Inner London Outer London Rest of South East West Midlands (Metropolitan) Strathclyde

First-degree Price Discrimination and Quality Customisation Under Data Protection Regulations

Tao Chen

University College London

Abstract

In response to privacy and ethical concerns, data protection laws such as the General Data Protection Regulations (GDPR) have now been put in place and ought to have an impact on the industries that are closely associated with price and quality customisation. Consumers now have a say in their personal data and can legally opt out of data-oriented personalisation schemes at their discretion. In this paper, I develop a Hotelling-styled spatial model to explore the interaction between the regulations and the industry in a duopolistic setting. In different scenarios, I show such legally binding options to opt out might either not increase consumer surplus or increase consumer surplus at the cost of social welfare.

**Thanks are due to Professor Martin Cripps, Dr Nikita Roketskiy, without whose advice this work would not be possible. I'm also very grateful for the continuous support from Dr Frank Witte and Dr Dunli Li.*

Table of Contents

1 Introduction

Tailored price and quality are becoming a new norm amid rapid growth in so-called 'database relationship marketing' thanks to the booming digital market and increasingly connected internet. The general public keeps updating the realisation of the profound impact by mass processing and analysis of personal data on the society. Shiller (2014[**shiller2014first**]) shows that an online service provider such as Netflix could earn more profit using web-browsing variables, compared to just using demographics to price their products. This is just one of the examples[1] that shows how this data could reveal the types and willingness to pay of the consumers for the benefits of the firms. News like data breaches or unethical use of data often worries society and inflates concerns over data privacy, which has led to the introduction of data protection laws in many parts of the world. This attracts inquiries regarding the economic impacts on the market and social welfare. This dissertation aims to provide an economic perspective on the debates over this legislation and contributes to the studies of price and product personalisation.

In Section Two, I discuss the related literature, especially those concerning competitive first-degree price discrimination and quality personalisation. I also provide an overview of the new data protection regulation and how it applies to this context. In Section Three, I set up the model and discuss five subgames respectively. In Sections Four and Five, I provide discussion and conclusions.

2 Related Literature

2.1 Price Discrimination

Stigler (1978[**stigler**]) defined price discrimination (PD) as

1 For a detailed survey on price personalisation in real world practice, refer to Borgesius and Poort,2017[**borgesius2017online**]

'the sale of two or more similar goods at prices that are in different ratios to marginal cost.' The theoretical literature of industrial organisation in the past decades saw a variety of interpretations approaches taken when modelling price discrimination (PD). As categorised in a survey by Stole (2007 [**stole2007price**]), PD could be discussed in third-degree PD, second-degree PD (non-linear pricing), purchase-history PD, product bundling etc. respectively. The rise of PD is commonly considered as a product of imperfect competition and asymmetric information, and is conducted for the purpose of surplus extraction. However, the first-degree PD was lesser discussed because of its impracticality and infeasibility until this millennium. Particularly because the rapid development of the e-commerce sector gives prevalence to 'personalisation' and allows firms to collect personal data with ease (e.g. online cookies usually record the web browsing data and feed back to the retailers and advertisers). Together with the purchasing mode of e-commerce, that transactions happen in a relatively isolated manner instead of posted pricing, one of the most important conditions for the existence of PD is satisfied: consumers can be segmented either directly or indirectly. For instance, firms can provide different offers to consumers with different characteristics e.g. gender, race etc. In the case of first-degree PD, consumers must be segmented at the individual level, which is technically possible in the current state of technology.

Price discrimination is commonly discussed in different competitive environments: monopoly, oligopoly and monopolistic competition with free entry. The welfare effect of first-degree PD under monopoly is immediate: social welfare is maximised as firms appropriate all the consumer surplus possible. This kind of setting rarely exists in reality, although it often serves as a benchmark. There are also marketing studies that depart from the simplest setting exploring other forms of pricing. Rayna et. al. (2015[**rayna2015pricing**]) proposed a pricing protocol that could facilitate a 'mutually advantageous' first-degree price discrimination in the digital music market under monopoly, which makes perfect PD desirable with certain conditions. As for monopolistic competition with

free entry, Spulber (1979[**spulber1979non**]) found the existence of a unique pure-strategy Nash equilibrium and Bhaskar and To (2004[**bhaskar2004perfect**]) concluded perfect PD always causes excessive entry from a social perspective.

The majority of literature focuses on oligopoly as it is thought that firms must have some market power to adopt the generally-known-to-be costly personalised technology and have access to sufficiently rich data. Hotelling's linear city model and Salop's circle model often serve as the base of the analytical literature should competition be introduced. The framework also provides an opportunity to incorporate consumers' heterogeneity in types which could be interpreted as brand/location preference or willingness to pay, subject to the research question. Ulph and Vulkan (2000 [**ulph2000electronic**]) presented a model in duopolistic competition discussing first-degree PD. Their model is a location-model where two firms are located at both ends of the uniformly distributed consumers who buy fixed amounts of goods. Two firms compete in Bertrand-style in price until an equilibrium is reached where consumers are indifferent about buying from either firm. In this model the product differ- entiation is represented through the location, and consumers simply derive a higher surplus with lower price (after deducting transport cost). In other words, the individual characteristics of the consumers are their locations through which firms learn how much utility will be derived by the consumers. They found that whether it is more profitable for firms to use first-degree PD depends on how the transport cost function is structured. Therefore, the dominance of either 'intensified competition effect' or 'enhanced surplus effect' is ambiguous, which is consistent with much competitive price discrimination literature(Varian, 1989[**varian1989price**]; Armstrong, 2006[**armstrong-2006competition**]).

There are studies adopting Hotelling-type models that discuss the dynamics of PD. This is considered to play a crucial role, as personalised pricing is often associated with the consumers' behaviour called Behaviour-based price discrimination (BBPD). This is where there has to be a stage in which consumers' purchases are being observed (i.e. information

acquisition) such that personalisation can be enabled in the later stage. Hence firms have an incentive to seize large first stage market shares with aggressive pricing such that they can exploit the advantages over the information they collect from the 'initial purchase.' Choe et. al. (2018 [**choe2017pricing**]) developed a dynamic model encompassing perfect targeting based on framework of Fudenberg and Tirole (2000 [**fudenberg2000customer**]) where third-degree price discrimination is used through the information firms gathers about market segmentation of loyalty. Their model predicts two asymmetric equilibria, and the personalised pricing strategy leads to a prisoners dilemma for firms with profits lowered compared to uniform pricing, in contrast to the symmetric equilibrium shown in the Fudenberg and Tirole model. For the dynamic models[2] with symmetric set-up of firms, they are consistent in that BBPD hurts profits by intensifying competition.

The work mentioned so far is done in a one-stop shopping[3] set-up of horizontal or vertical product differentiation for a single product. Horizontal product differentiation is the difference in brand preference of the consumer represented through dimensional location advantage. While with vertical product differentiation, firms rank the consumer types identically hence consumer's taste of quality is independent of brand preference. Single-dimensional models incur inadequacy to capture both brand preferences and the marginal value of consumption. The multi-dimensional model is approached either through simulation (Borenstein and Rose, 1994[**borenstein1994competition**]) or discrete-choice approach (Mussa and Rosen, 1978[**mussa1978monopoly**]; Rochet and Stole, 2002[rochet2002nonlinear]). Ghose et. al. (2009[**ghose2009personalized**]) manage to capture both quality and brand preference by making a stylised assumption

2 Chen (1997[**chen1997paying**]); Villas-Boas (1999[**villas-1999dynamic**]); Fudenberg and Tirole(2000[**esteves2010pricing**]); Pazgal and Soberman (2008[**pazgal2008behavior**])

3 In one-stop shopping, consumers may desire at most one product but have preference on qualityand they choose to purchase from the firms with highest indirect utility.

when specifying their model; the location of the consumer simultaneously denotes both preferences. In other words, it is assumed that the larger the brand preference of a given consumer for one firm, the larger is the quality preference of that consumer for that firm's product, which is well backed by empirical studies and intuitions.

In the setup of Ghose et. al., (2009), it is convenient to investigate both person- alised pricing and quality customisation, contrasting with the majority of the other models. This is an important feature, as it is closely associated with one of the most common practices in real world marketing known as versioning. Recall the definition of price discrimination by Stigler: 'two similar goods' instead of two identical goods. Firms usually have a lineup of products that fall into the same cluster on the spectrum of product differentiation in the whole market, but still differentiate in feature and performance for different target groups. For instance, Apple annually releases two or three new generation iPhones that are similar but differentiate in subtle features. The premier version is typically sold with a much higher margin. It has also seen a prevailing trend that consumers are in favour of tailored, personalised product and service, which provides room for firms to achieve product customisation at its extreme form with data analytic technology. This means personal data processing is not only concerning the price but also the quality and that neither should be omitted from analysis. Therefore, the modelling of this work is based on the framework of Ghose et. al. which is also benchmarked. Details of the model are deferred to later sections where I introduce the set-up of the model.

2.2 *Fairness, Data Privacy, and Regulations*

The discussion of personalising technology and price discrimination has never been restricted to the field of economics. In a survey by Kahneman et. al. (1986[**kahneman-1986fairness**]), 91% of respondents viewed personalised pricing as unfair. How firms utilise the marketing technology without triggering backlash has been a real challenge. (Tan-

ner, 2014[**tanner2014different**]). Now being viewed as a classic case study, Amazon once experimented with personalised pricing but had to retract after the move backfired. Without going further into case studies and surveys, it is obvious that people are generally uncomfortable with personalised pricing. Besides what is suggested in behavioural economics such as regret aversion(Loomes and Sugden, 1982[**loomes1982regret**]), consumers feel it is unfair because they are paying different prices for products that are exactly the same as in the Amazon case. However, their psychology changes when firms conduct versioning as aforementioned; the extra premium on high quality product could be easily justified and accepted. This means personalisation marketing might not necessarily backfire when product quality is customised as well.

Some information theoretic literature discusses the trade-off between privacy loss and profits, which effectively concern the optimal pricing mechanism design. Those studies are motivated by the different valuation of private information; firms usually extract more rent with information from certain types of consumers (e.g. high types). Hence firms don't necessarily have an incentive to acquire all the information that covers the whole market when information collection is costly. Eliat et. al.(2019[**eliaz2019optimal**]) propose a Bayesian measure of loss of privacy to incorporate privacy constraints into mechanism design. The strategic interaction between the pricing mechanism and the consumer sparks many interesting questions with regard to those that are often omitted or endogenously assumed away in the industrial organisation literature. However, it's beyond the scope of this context for the focus of the research question. Therefore the model introduced in the next section is still based on commonly stylised assumptions such as perfect information and zero information acquisition cost for the sake of simplicity.

In the wake of the 'Cambridge Analytica' scandal in 2018, European legislators quickly responded by upgrading data regulations. The regulations on the usage of personal data are the most stringent in history. Under General Data Protection

Regulation (GDPR)[4], the binding law across European Union and European Economic Area, consumer consent is required for firms to process personal data, and consumers can 'opt out' of the agreement anytime by discretion. Scholars in legal studies commonly agree that GDPR applies to personalised pricing in general and cookies with unique identifiers should be regarded as personal data(Borgesius and Poort, 2017[**borgesius20170online**]; Steppe, 2017[**steppe20170online**]). Therefore, consumers are legally granted more power for them to play a more important role in the game of personalisation. Some studies incorporate the strategic behaviours of consumers for example by delaying consumption in the early period to gain an advantage in the future (Chen and Zhang, 2009[**chen2009dynamic**]). The ability to opt out of personalisation programmes should have more robust implications, which is the prime motivation of this research.

3 The Model

3.1 The Set-up

Here we construct a duopoly model with two multi-product firms which set price and quality to compete. The model is based on the framework of Ghose et. al.(2009) (henceforth G09) where each firm's product line consists of a continuum of qualities and is fixed in length. Consumers types are denoted as θ where $\theta \in [0, 1]$ with uniform distribution. This draws the uniqueness of G09 as θ represents both the brand preference and the marginal valuation of quality. This set-up implicitly assumes that consumers with stronger brand preference (located closer to the end) would derive a higher utility from an increase in quality supplied by the preferred firm and, hence would have a higher willingness to pay (WTP) for the quality supplied. The interpretation provides a nice way around the dichotomy in horizontal and vertical product differentiation and is well supported by empirical studies. G09 gave an example in the automobile industry: consumers who prefer brands specialised in highly fuel-efficient vehicles

4 eugdpr.org

would place more value on fuel efficiency and less value on other dimensions such as outdoor performance valued by SUV lovers. (Goldberg, 1995[**goldberg1995product**]; Berry et. al., 1995[**berry1995automobile**]). Other examples include the enterprise software industry and commercial airplane industry where products are often highly customised as prices are customer-specific, based on negotiations.

In a standard Hotelling-style setup, two firms, call them firm L and firm R, are located on the two ends of the 'linear city' where consumers are uniformly distributed and ranked by type. Quasi-linear utility function is assumed: the gross utility to a consumer with type θ buying from the firm located at 0, firm L is

(1) $$u^L(q, \theta) = q \times (1 - \theta)$$

his gross utility derived from buying from the firm located at 1, firm R, is

(2) $$u^R(q, \theta) = q \times \theta$$

the firm sets a price according to consumer types (targeted or not) pL(θ), pR(θ), the indirect utility/consumer surplus function of the consumers are

(3) $$s^L(q, \theta) = q \times (1 - \theta) - p^L(\theta)$$

(4) $$s^R(q, \theta) = q \times \theta - p^R(\theta)$$

identical to G09, assuming the marginal cost of production is invariant with quantity but depends on the quality of the product. Depending on the quality schedule, both firms have the identical cost function

(5) $$c(q) = q^\alpha/\alpha,\ \alpha > 1$$

The sequence of the game is as follows: In the first period, the firms simultaneously choose the pricing strategy they will use for two pools of consumers: those who opt in and

opt out[5] of the personalisation programme. Firms are neither able to process the data of out-consumers nor use their data against them by perfect targeting. In other words, once consumers opt out, their location becomes private information and targeting is disabled, which means they won't get the product with price and quality that the firms personalise just for them. Therefore, personalised price and quality (PPQ) is ruled out for the out-consumers by definition. In this model, we consider three pricing strategies: personalised price and quality (PPQ); non-targeting non-linear pricing (NNP); uniform pricing. PPQ is the firm's offering of a pair of price and quality designed specifically for some consumer type; the consumer either accepts the offer or does not purchase. NNP is a classic form of second-degree PD that relies on the consumers self-selection; firms set a price and quality schedules for each type of consumers, and consumers self-select them into buying the quality-price pair that maximises their utility from the menu. NNP is subject to consumer incentive compatibility and individual rationality constraints such that he purchases the product with price and quality of his type. Uniform pricing is when firm doesn't discriminate at all and offer the same quality and price for every consumer. Therefore there are those cases of combination of firms' choice to consider:

I Both firms use PPQ for in-consumers and NNP for out-consumers

II Both firms use PPQ for in-consumers and uniform pricing for out-consumers

III Both firm use PPQ for in-consumers; one firm uses NNP for out-consumers, onefirm uses uniform pricing for out-consumers

IV one firm uses PPQ for in-consumers and NNP for out-consumers, the other firmuses NNP for all consumers

V one firm uses PPQ for in-consumers and NNP for out-con-

5 henceforth in-consumers and out-consumers.

sumers, the other firm uses uniform pricing for all consumers

Note the latter two cases consider the scenarios where only one firm has the access to personalised technology or forego the technology. Because the firms are symmetric hence the analysis can depart from one firm and the conclusion still holds for the other firm. As the data protection regulations allow consumers to pull out any time and firms are obliged to comply, we can assume that consumers make the decision after observing the pricing strategy by the firms. This also assumes that consumers have the price and quality information in both segments hence prefers the segment in which they get higher surplus. The game then precedes as follows. Consumers make the decision of whether to opt into personalisation programme, firms then set price-quality accordingly for thetwo segments, the in-consumer segment or out-consumer segment. An equilibrium isreached when there is no profitable deviation between both firms and consumers. Noteit's assumed that firms do not switch pricing strategies as it complicates analysis.

3.2 Subgame I: both firms use PPQ for in-consumer, NNP pricing for out-consumer

From G09, we learn that the consumers at the middle always gain the highest surplus from PPQ while the most loyal consumers (those locate close to 1 or 0) get the lowest with their surplus fully extracted. Hence, the in-consumer segment will be continuousin the middle thanks to the quasi-linearity of the utility function. Suppose for consumer type $\theta \in [\hat{\theta}, 1-\hat{\theta}]$ choose to option and otherwise opt out. Denote $A=\hat{\theta}$; $C=1-\hat{\theta}$ such that $A+C=1$ with A and C being the marginal consumers on the edge of opting out. Denote the marginal consumer who feels indifferent between the personalised offerby both firms by $\theta=B$. In this subgame, each firm offers a menu i.e. quality-price schedule for out consumers $\theta \in [0, A)$; $\theta \in (C, 1]$ and a pair of personalised price and quality for in-consumers $\theta \in [A, C]$. The decision variables of the firms are $q(\theta)$ and $p(\theta)$ for all segment. Focusing the analysis on firm R, the

objective function is given by

$$\max_{p^R(\theta), q^R(\theta)} \pi^R, \quad \text{where} \quad \pi^R = \int_B^1 \left[p^R(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha} \right] d\theta \tag{6}$$

Starting from the out-consumer segment, standard in self-selection literature, we need an incentive compatibility (IC) constraint to ensure that consumer purchases the product designed for his type instead of other types such that his utility is maximised from the purchase. Additionally, the consumer should derive non-negative indirect utility for him to accept any price-quality bundle at all, which forms an individual rationality (IR) constraint. In addition to participation condition, firms have to offer consumer surplus no less than that offered by the rival to keep the consumer. Therefore the optimisation problem is subject to the following constraints:

- **IC:** $\theta = \arg\max_t \quad \theta \times q^R(t) - p^R(t), \forall \theta \in [C, 1]$ where t denotes the type of the chosen product.
- **IR(1):** $s^R(\theta) \geq 0, \forall \theta \in [C, 1]$
- **IR(2):** $s^R(\theta) \geq s^L(\theta)$

Note the reason why the domain of the constraints are restricted to [C,1] is not the constraints only apply for the consumers of the segment: all consumers are still free to choose between two firms in this model. Because of the specification of the utility function, it is easy to show that in equilibrium compared with marginal consumer C, the consumers to the left of C, those in the out-consumer segment loyal to firm L strictly prefer buying from firm L, and vice versa. Therefore, firm R serves [C, 1] and firm L serves [0, A]. This is inherited from Mussa and Rosen (1987[**mussa1978monopoly**]) and G09.

By (4) transforming the decision variable the price schedule $p^R(\theta)$ to a function of $s^R(\theta)$, we have $p^R(\theta) = \theta q^R(\theta) - s^R(\theta)$ and

rewrite the objective function of firm R as:

(7) $$\max_{s^R(\theta),q^R(\theta)} \pi^R, \quad \text{where} \quad \pi^R = \int_C^1 \left[\theta q^R(\theta) - s^R(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha}\right] d\theta$$

The objective function for firm L by symmetry can be written as following:

(8) $$\max_{s^L(\theta),q^L(\theta)} \pi^L, \quad \text{where} \quad \pi^L = \int_0^A \left[(1-\theta)q^L(\theta) - s^L(\theta) - \frac{(q^L)^\alpha(\theta)}{\alpha}\right] d\theta$$

As per IC constraint, consumer gains the indirect utility:

(9) $$S^R(\theta) = \max_t \quad \theta \times q^R(t) - p^R(t)$$

The first order condition is:

(10) $$\theta \times \frac{\partial q^R(t)}{\partial t} - \frac{\partial p^R(t)}{\partial t} = 0$$

Using the envelop theorem, the equation holds at t=θ because the consumer self-select the the price and quality pair designed for his type. By differentiating 9:

(11) $$\frac{ds^R(\theta)}{d\theta} = q^R(\theta) + \theta \frac{\partial q^R(\theta)}{\partial \theta} - \frac{\partial p^R(\theta)}{\partial \theta}$$

Hence we have:

Lemma 1.

(12) $$\frac{ds^R(\theta)}{d\theta} = q^R(\theta)$$

(13) $$\frac{ds^L(\theta)}{d\theta} = q^L(\theta)$$

This implies that the quality schedule $q^R(\theta)$ equals to the slope of the consumer surplus $s^R(\theta)$ and we obtain that

(14) $$s^R(\theta) = s^R(C) + \int_C^\theta q^R(t)dt$$

(15) $$s^L(\theta) = s^L(A) + \int_\theta^A q^L(t)dt$$

In G09, firms compete on the marginal consumer who is indifferent between buying from two firms. Hence there is an additional IR constraint to ensure that $s^R(B) = s^L(B)$ and the firms thus compete by lowering the pricing schedule by a constant, $s^R(B)$. Asa result, loyal consumers receive higher surplus termed as information rent. However,in the setting of this model, the existence of the in-consumer segment in the middleground forms a barrier to facilitate two local monopoly: it's straightforward that for the marginal consumers of two segments $A \neq C$, they strictly prefer the offer by the closer firm as it is not profitable for the further firm to poach the marginal consumer of the loyal out-consumer segment of the rivalry due to location disadvantage. Therefore,similar to a monopoly, firms drive the surplus of the marginal consumer to zero such that $s^L(A) = s^R(C) = 0$. The optimisation problem can be rewritten as:

(16) $$\max_{\{s^R(\theta), q^R(\theta)\}} \pi^R, \quad \text{where} \quad \pi^R = \int_C^1 \left[\theta q^R(\theta) - s^R(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha}\right] d\theta$$

(17) $$s.t. s^R(\theta) \geq 0, s^R(C) = s^L(A) = 0$$

substituting for $s^R(\theta)$, we have

(18) $$\max_{s^R(\theta)} \pi^R, \quad \text{where} \quad \pi^R = \int_C^1 \left[\theta q^R(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha} - \int_C^\theta q^R(t) dt\right] d\theta$$

Changing the integration order of last term in the bracket:

(19) $$\int_C^1 [\int_C^\theta q^R(t) dt] d\theta = \int_c^1 [\int_C^1 q^R(t) d\theta] dt = \int_C^1 q^R(t)(1-t) dt = \int_C^1 q^R(\theta)(1-\theta) d\theta$$

then the objective function becomes

(20) $$\max_{q^R} \pi^R \text{where} \quad \pi^R = \int_C^1 \left[\theta q^R(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha} - q^R(\theta)(1-\theta)\right] d\theta$$
$$= \int_C^1 \left[(2\theta - 1) q^R(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha}\right] d\theta$$

Similarly, the optimisation problem for firm L follows:

(21) $$\max_{q^L(\theta)} \pi^L \text{where} \quad \pi^L = \int_0^A \left[(1-2\theta)q^L(\theta) - \frac{(q^L)^\alpha(\theta)}{\alpha}\right] d\theta$$

To obtain the quality schedule, differentiate the terms in (10) with respect to $q^R(\theta)$:

$$\theta - (q^R)^{\alpha-1}(\theta) - (1-\theta) = 0$$

$$\Rightarrow$$

Lemma 2
The quality schedules are
$q^R(\theta) = (2\theta - 1)^{\frac{1}{\alpha-1}}$ *and* $q^L(\theta) = (1-2\theta)^{\frac{1}{\alpha-1}}$.

Now we can obtain the consumer surplus function by Lemma 1 and Lemma 2:

(22) $$s^L(\theta) = 0 + \int_\theta^A q^L(t)dt$$

(23) $$s^L(\theta) = 0 + \int_C^\theta q^R(t)dt$$

We can obtain the price schedules by $p^L(\theta) = (1-\theta)q^L(\theta) - s^L(\theta)$ and $p^R(\theta) = \theta q^R(\theta) - s^R(\theta)$. Therefore we have the optimal price schedule:

(24) $$p^L(\theta) = (1-\theta)(1-2\theta)^{\frac{1}{\alpha-1}} - \frac{\alpha-1}{2\alpha}(1-2\theta)^{\frac{\alpha}{\alpha-1}} + \frac{\alpha-1}{2\alpha}(1-2A)^{\frac{\alpha}{\alpha-1}}$$

(25) $$p^R(\theta) = \theta(2\theta-1)^{\frac{1}{\alpha-1}} - \frac{\alpha}{\alpha-1}(2\theta-1)^{\frac{\alpha}{\alpha-1}} + \frac{\alpha-1}{2\alpha}(2B-1)^{\frac{\alpha}{\alpha-1}}$$

The first two terms are identical to G09 in the case of both firms conducting NNP. The NNP price is strictly higher in the presence of the in-consumer segment since the last term is strictly positive. The increase in price should fully extract the consumer surplus of the marginal consumer at the edge of out-segment. Now consider the in-consumer segment, the analysis will be less different from the G09 since the segment is just a truncated market as in their original model. Unlike the out-consumer segments, the in-consumer segment will be a continuum of consumers at the middle and firms face

direct competition from each other: there may be incentive to poach consumers for market share. Recall the objective functions of two firms:

(26) $$\max_{s^R(\theta), q^R(\theta)} \pi^R, \quad \text{where} \quad \pi^R = \int_C^1 \left[\theta q^R(\theta) - s^R(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha}\right] d\theta$$

(27) $$\max_{s^L(\theta), q^L(\theta)} \pi^L, \quad \text{where} \quad \pi^L = \int_0^A \left[(1-\theta) q^L(\theta) - s^L(\theta) - \frac{(q^L)^\alpha(\theta)}{\alpha}\right] d\theta$$

Due to perfect targeting by PPQ, there is no self-selection concern here hence we can make IC constraint redundant. Therefore $s^L(\theta)$ and $s^R(\theta)$ are equal to the socially optimal surplus. This contrasts the quality degradation due to the fear of product cannibalisation seen in NNP case.

First order condition gives the quality schedule:

(28) $$\frac{\partial \pi^L(\theta)}{\partial q^L(\theta)} = (1-\theta) - (q^R)^{\alpha-1}(\theta) = 0$$
$$\Leftrightarrow q^L(\theta) = (1-\theta)^{\frac{1}{\alpha-1}}$$

(29) $$\frac{\partial \pi^R(\theta)}{\partial q^R(\theta)} = \theta - (q^R)^{\alpha-1}(\theta) = 0$$
$$\Leftrightarrow q^R(\theta) = \theta^{\frac{1}{\alpha-1}}$$

As firms have the perfect information about the in-consumers, they will compete in a Bertrand manner at the individual level, which means the closer firm will offer exactly the highest possible surplus offered by the other firm to keep the consumer. Each consumers will get offers from both firms and should feel indifferent from accepting either one of them. The closer firm appropriates the remaining surplus. Thus the consumer surplus $s^L(\theta)$ and $s^R(\theta)$ should equal to the socially optimal surplus by the rival firm. The rival firm offers socially optimal quality that maximise its profit and marginal cost (due to Bertrand price competition) to the consumers located closer to the other firm. We obtain

(30) $$s^L(\theta) = \max_{q^R(\theta)} \left[\theta q^R(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha}\right] = \left(1 - \frac{1}{\alpha}\right)\theta^{\frac{\alpha}{\alpha-1}}, \theta \in [A, 1/2]$$

(31) $$s^R(\theta) = \max_{q^L(\theta)} \left[(1-\theta)q^L(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha}\right] = \left(1 - \frac{1}{\alpha}\right)(1-\theta)^{\frac{\alpha}{\alpha-1}}, \theta \in [1/2, C]$$

In equilibrium, all consumers in [A, 1/2] buy from firm L and all consumers in [1/2, C] buy from firm R. Substituting in optimal quality schedule, we have the consumer surplus

(32) $$S^L(\theta) = (1 - \frac{1}{\alpha})\theta^{\frac{\alpha}{\alpha-1}}$$

(33) $$S^R(\theta) = (1 - \frac{1}{\alpha})(1-\theta)^{\frac{\alpha}{\alpha-1}}$$

By p(θ) =u(q(θ),θ)–s(θ) we have the price schedule for in-consumer segment:

(34) $$p^L(\theta) = (1-\theta)^{\frac{\alpha}{\alpha-1}} - (1 - \frac{1}{\alpha})\theta^{\frac{\alpha}{\alpha-1}}, \theta \in [A, 1/2]$$

(35) $$p^R(\theta) = \theta^{\frac{\alpha}{\alpha-1}} - (1 - \frac{1}{\alpha})(1-\theta)^{\frac{\alpha}{\alpha-1}}, \theta \in [1/2, C]$$

Now we have the quality and price schedule for both segments as we arrive at

Lemma 3 *Given the size of the segments, when both firms use PPQ for in-consumers and NNP for out-consumers, the best response prices, quality schedules and surplus function are the following:*

$$q^L(\theta) = \begin{cases} (1-2\theta)^{1/(\alpha-1)} \text{if} & \theta \in [0, A] \\ (1-\theta)^{1/(\alpha-1)} \text{if} & \theta \in [A, C] \\ 0 \quad \text{if} & \theta \in [C, 0] \end{cases}$$

$$q^R(\theta) = \begin{cases} (2\theta-1)^{1/(\alpha-1)} \text{if} & \theta \in [C, 1] \\ \theta^{1/(\alpha-1)} \text{if} & \theta \in [A, C] \\ 0 \quad \text{if} & \theta \in [0, A] \end{cases}$$

$$s^L(\theta) = \begin{cases} \frac{\alpha-1}{2\alpha}(1-2\theta)^{\alpha/(\alpha-1)} - \frac{\alpha-1}{2\alpha}(1-2A)^{\alpha/(\alpha-1)} \text{if} & \theta \in [0, A] \\ (1-\frac{1}{\alpha})\theta^{\alpha/(\alpha-1)} & \text{if} \quad \theta \in [A, 1/2] \end{cases}$$

$$s^R(\theta) = \begin{cases} \frac{\alpha-1}{2\alpha}(2\theta-1)^{\alpha/(\alpha-1)} - \frac{\alpha-1}{2\alpha}(2C-1)^{\alpha/(\alpha-1)} \text{if} & \theta \in [C, 1] \\ (1-\frac{1}{\alpha})(1-\theta)^{\alpha/(\alpha-1)} & \text{if} \quad \theta \in [1/2, C] \end{cases}$$

$$p^L(\theta) = \begin{cases} (1-\theta)(1-2\theta)^{\frac{1}{\alpha-1}} - \frac{\alpha-1}{2\alpha}(1-2\theta)^{\frac{\alpha}{\alpha-1}} + \frac{\alpha-1}{2\alpha}(1-2A)^{\frac{\alpha}{\alpha-1}} \text{if} & \theta \in [0, A] \\ (1-\theta)^{\alpha/(\alpha-1)} - (1-\frac{1}{\alpha})\theta^{\alpha/(\alpha-1)} & \text{if} \quad \theta \in [A, 1/2] \\ 0 & \text{if} \quad \theta \in [1/2, 1] \end{cases}$$

$$p^R(\theta) = \begin{cases} \theta(2\theta-1)^{\frac{1}{\alpha-1}} - \frac{\alpha}{\alpha-1}(2\theta-1)^{\frac{\alpha}{\alpha-1}} + \frac{\alpha-1}{2\alpha}(2B-1)^{\frac{\alpha}{\alpha-1}} \text{if} & \theta \in [C, 1] \\ p^R(\theta) = \theta^{\frac{\alpha}{\alpha-1}} - (1-\frac{1}{\alpha})(1-\theta)^{\frac{\alpha}{\alpha-1}} & \text{if} \quad \theta \in [1/2, C] \\ 0 & \text{if} \quad \theta \in [0, 1/2] \end{cases}$$

The original model of G09 considers the scenario for all consumers where both firms engage in PPQ and both firms engage in NNP, which can be considered as a special case with all consumers opt in and all consumers opt out: *A*=*C*= *1/2*. We canalso consider G09 as a state of world where data protection regulation have not been introduced yet as this can serve as the first round of the game. Hence, after firms decide on the pricing strategies, in round 0, they set price-quality for in-consumers as if all consumers are in the personalisation programme and price-quality schedulesfor out-consumers as if all consumers opt out. We know that the consumer in the middle enjoys the highest surplus while the most loyal consumer getting surplus fullyextracted when both firms use PPQ. The reverse is true when both firms use NNP. This is because in PPQ the least loyal consumers (the ones at the middle) benefit from the fierce competition between firms drive price down to prevent theirs customers from being poached. Under NNP, the most loyal consumers enjoy 'information rent' (Mussaand Rosen, 1978[**mussa1978monopoly**]) due to firm's fear of product

cannibalisation.To find the marginal consumer ˆθ that is indifferent between opting in/out, equate the consumer surplus (purchasing from firm L) for both pricing and solve forˆθ:

(1)
$$\frac{\alpha-1}{2\alpha}(1-2\hat{\theta})^{\alpha/(\alpha-1)} = (1-\frac{1}{\alpha})\hat{\theta}^{\alpha/(\alpha}$$
$$\Rightarrow \hat{\theta} = \frac{2^{(1-\alpha)/\alpha}}{1+2^{1/\alpha}}$$

By symmetry,

(2)
$$1-\hat{\theta} = 1-\frac{2^{(1-\alpha)/\alpha}}{1+2^{1/\alpha}}$$

Therefore in round 0, the consumers with type $\theta \in \left[0, \frac{2^{(1-\alpha)/\alpha}}{1+2^{1/\alpha}}\right]$ and $\left[1-\frac{2^{(1-\alpha)/\alpha}}{1+2^{1/\alpha}}, 1\right]$ the loyal types, want to opt out to receive higher surplus while the less loyal types prefer to opt in. Therefore this forms a first round segmentation: consumer type $\theta \in \left[\frac{2^{(1-\alpha)/\alpha}}{1+2^{1/\alpha}}, 1-\frac{2^{(1-\alpha)/\alpha}}{1+2^{1/\alpha}}\right]$ choose to opt in and otherwise opt out. Subsequently, firms update their price-quality menu for the out-consumer segment. Since the competitive nature hasn't changed, firms continue to offer the same PPQ price and quality as before when there is no out-segment. From the derivation before, we know that the quality schedule follows the IC constraint and is not determined by the size of the segment. Hence price schedule effects all the rent extraction at this stage. Recall the price schedule for out-consumers:

(3)
$$p^L(\theta) = (1-\theta)(1-2\theta)^{\frac{1}{\alpha-1}} - \frac{\alpha-1}{2\alpha}(1-2\theta)^{\frac{\alpha}{\alpha-1}} + \frac{\alpha-1}{2\alpha}(1-2A)^{\frac{\alpha}{\alpha-1}}$$

(4)
$$p^R(\theta) = \theta(2\theta-1)^{\frac{1}{\alpha-1}} - \frac{\alpha}{\alpha-1}(2\theta-1)^{\frac{\alpha}{\alpha-1}} + \frac{\alpha-1}{2\alpha}(2B-1)^{\frac{\alpha}{\alpha-1}}$$

The size of increase in price, $\frac{\alpha-1}{2\alpha}(1-2A)^{\frac{\alpha}{\alpha-1}}$ $\frac{\alpha-1}{2\alpha}(2B-1)^{\frac{\alpha}{\alpha-1}}$ depends on the size of the out-segment indicated by A,C and should appropriate all the consumer surplus for those at the margin. Plugging in $A = \hat{\theta} = \frac{2^{(1-\alpha)/\alpha}}{1+2^{1/\alpha}}$ and

$C = 1 - \hat{\theta} = 1 - \frac{2^{(1-\alpha)/\alpha}}{1+2^{1/\alpha}}$ we can obtain the updated price schedule for out-segment. In the next round, it's straightforward to show that some consumers who have opted out in the last round is worse off due to the new price schedule and would like to opt in. The in-segment is hence larger and firms update the price schedule accordingly. As shown in Figure 1[6],this iteration process keeps going until only the most loyal consumer, $\theta = 0$ or $\theta = 1$ is left in the out-segment. At that time, they should be indifferent to opting is and opting out as their surplus is fully extracted either way. Therefore we arrive at:

Proposition 1. *When both firms use PPQ for in-consumers and NNP for out-consumers,all consumers choose to opt in and receive PPQ offer in equilibrium.*

Intuitively, because firms hold the location information of the in-consumers and hence the size of the segment, firms do not directly need the consumer to opt in for them to know the type of the consumers at the margin of out-segment. The information is crucial as firms can accordingly extract maximum surplus by updating price schedule and there is no fear of rival firm poaching customers. In this subgame, even though consumers are given the discretionary choice of staying in and out, the typical consumer is weakly worse off when opting out, that is, the utility-maximising product from the quality-price menu for the out-segment delivers no greater surplus comparing to the offer that is customised for the consumer. From the perspective of social welfare, it's straightforward to show that the social welfare is highest in equilibrium and it's socially optimal. Firms offer sub-optimal quality schedule, a.k.a. product degradation under NNP to prevent product cannibalisation under NNP but offer socially optimal quality with PPQ. More on welfare discussion shall be discussed in more detail in the later section.

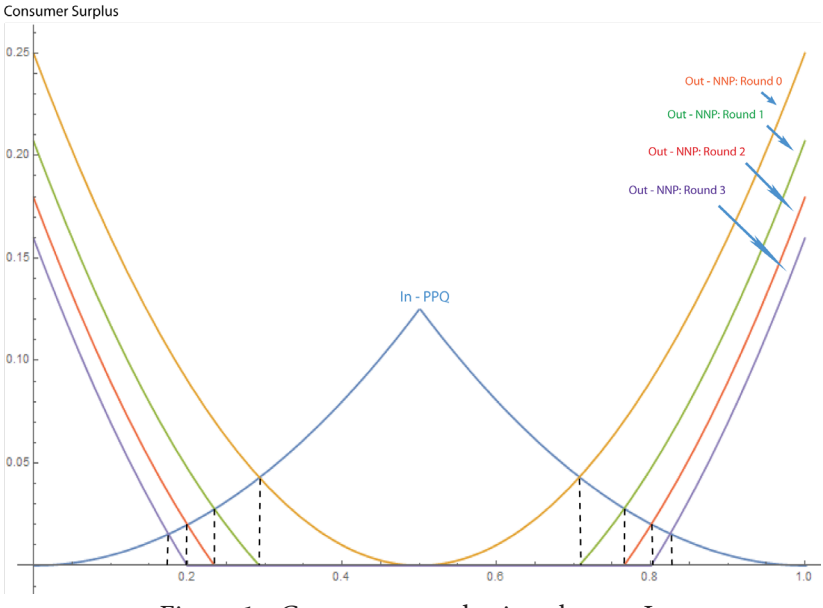


Figure 1 - Consumer surplus in subgame I

3.3 Subgame case II: both firms use PPQ for in-consumers, uniform pricing for out-consumers

The subsection analyses the subgame when both firms use PPQ for in-consumers and uniform pricing for out-consumers. The analysis for in-consumer segment will remain the same as in previous section for same pricing being used. For out-consumer segment, firms will offer a flat quality and price for every consumers with no regards to their types in this case. Firms choose a price and quality that maximises profits. Consumers choose to purchase from the firm with which they receive surplus equal or higher than 0. This implies that the consumer who is indifferent between buying and not buying, i.e. $s(\theta) = 0$ determines the optimal market share for firms: consumers will not purchase from the firm unless he is closer (or equally close) to the firm than the indifferent

consumer. Hence instead of considering firms setting price and quality to maximise profit, we can consider firms choosing the optimal market share to serve with quality that maximises profits. Without loss of generality, we again depart from the perspective of firm R and the analysis for firm L should follow by symmetry. We also start with the stage when no one has opted in. As the indifferent consumer (denoted by θ^*_R) derives zero surplus, from $s(\theta) = \theta q - p$, we have $p = \theta^* q$. We have the optimisation problem for firm R:

(5) $$\max_{\theta^*_R, q_R} \pi^R \quad \text{where} \quad \pi^R = (1-\theta^*_R)(\theta^*_R q_R - \frac{q_R^\alpha}{\alpha})$$

the optimisation problem for firm L:

(6) $$\max_{\theta^*_L, q_L} \pi^L \quad \text{where} \quad \pi^L = \theta^*_L((1-\theta^*_L)q_L - \frac{q_L^\alpha}{\alpha})$$

The first order conditions for firm R are:

(7) $$\frac{\partial \pi^R}{\partial \theta^*_R} = q_R - 2\theta^*_R q_R + \frac{q_R^\alpha}{\alpha} = 0$$

and that for firm L are

(8) $$\frac{\partial \pi^L}{\partial q_R} = \theta^*_R - q^{\alpha-1} - \theta^{*2} + \theta^*_R q^{\alpha-1} = 0$$

Solving for θ^*_R and q_R, we have

(9) $$\theta^*_R = \frac{\alpha}{2\alpha - 1}$$

(10) $$q_R = \left(\frac{\alpha}{2\alpha-1}\right)^{\frac{1}{1-\alpha}}$$

Then we obtain

Lemma 4. *When all consumers opt out, the best response uniform price, quality,consumer surplus set by the firms are the following:*

(11) $$p_L = p_R = \left(\frac{\alpha}{2\alpha - 1}\right)^{\frac{2-\alpha}{1-\alpha}}$$

(12) $$q_L = q_R = \left(\frac{\alpha}{2\alpha - 1}\right)^{\frac{1}{1-\alpha}}$$

$$s^R(\theta) = \begin{cases} \theta\left(\frac{\alpha}{2\alpha - 1}\right)^{\frac{1}{1-\alpha}} - \left(\frac{\alpha}{2\alpha - 1}\right)^{\frac{2-\alpha}{1-\alpha}} \text{if} \quad \theta \in \left[\frac{\alpha}{2\alpha - 1}, 1\right] \\ 0 \qquad \text{Otherwise} \end{cases}$$

$$s^L(\theta) = \begin{cases} (1-\theta)\left(\frac{\alpha}{2\alpha - 1}\right)^{\frac{1}{1-\alpha}} - \left(\frac{\alpha}{2\alpha - 1}\right)^{\frac{2-\alpha}{1-\alpha}} \text{if} \quad \theta \in \left[0, \frac{\alpha - 1}{2\alpha - 1}\right] \\ 0 \qquad \text{Otherwise} \end{cases}$$

We can see that as $\alpha \to \infty$, $\frac{\alpha}{2\alpha-1} \to \frac{1}{2}$ and $\frac{\alpha}{2\alpha-1} < \frac{1}{2}$ for $\alpha > 1$ for α >1 Therefore, it's never optimal for the firm to cover more than half of the market and there will always be consumers not purchasing i.e. the market is not covered. Equilibrium test is easily passed as there is no profitable deviation for firms to seek larger market share from the rival's territory. To find the marginal consumer who is indifferent between opting in/out, i.e. receives same surplus from uniform price and quality and PPQ, equate the consumer surplus (purchasing from firm R) function under two pricing:

(1) $$(1 - \frac{1}{\alpha})(1-\theta)^{\frac{\alpha}{\alpha-1}} = \theta\left(\frac{\alpha}{2\alpha - 1}\right)^{\frac{1}{1-\alpha}} - \left(\frac{\alpha}{2\alpha - 1}\right)^{\frac{2-\alpha}{1-\alpha}}$$

The consumer type of θ that satisfies the above equation is the marginal consumer of the in-consumer segment. The consumers to the left of the marginal consumer either receive a lower surplus comparing to the surplus they would have received under PPQ or do not purchase under uniform pricing. Hence it's optimal for them to opt in and receive a higher surplus. Note that after the first round

of opting in, it's straightforward to show that the market size of the out-consumer segment shrinks such that the optimal market share is no longer attainable because some consumers who purchase under uniform pricing will opt in once given the option. Hence the corner solution arises: firms cover the whole out-segment and sets price that exactly extracts all the consumer surplus at the margin to maximise profits. Then the iteration process as shown in figure 2[7] follows the same manner as in case I and reaches an equilibrium when consumers locating at the end, $\theta = 0$, $\theta = 1$ are indifferent between the uniform pricing offer and *PPQ* offer. We arrive at

Proposition 2. *When both firms use PPQ for in-consumers, uniform pricing for out-consumers, all consumers opt in and receive PPQ offer in equilibrium.*

This case is more intuitive than the last one where both firms use *NNP* for out-segment. Because of the nature of uniform pricing, the firms will leave out some consumer such that the market is not covered. Then there is nothing stopping the consumers in the middle from opting into the personalisation scheme to achieve higher surplus comparing to none. Given firms always maximise profits subject to the size of the out-segment, recall the first order condition:

(2) $$\frac{\partial \pi^L}{\partial q_R} = \theta_R^* - q^{\alpha-1} - \theta^{*2} + \theta_R^* q^{\alpha-1} = 0$$

Solve for q_R, given the constraint $C=\theta$

(3) $$q_R = \left(\frac{C^2 - C}{C - 1}\right)^{\frac{1}{\alpha-1}}$$

We can see for $\alpha > 1$, it's easy to show that q_R strictly increases in C. Hence as the out-segment get smaller, the quality provided by the firms will increase along price until they reach the maximum that the most loyal consumers are willing to pay.

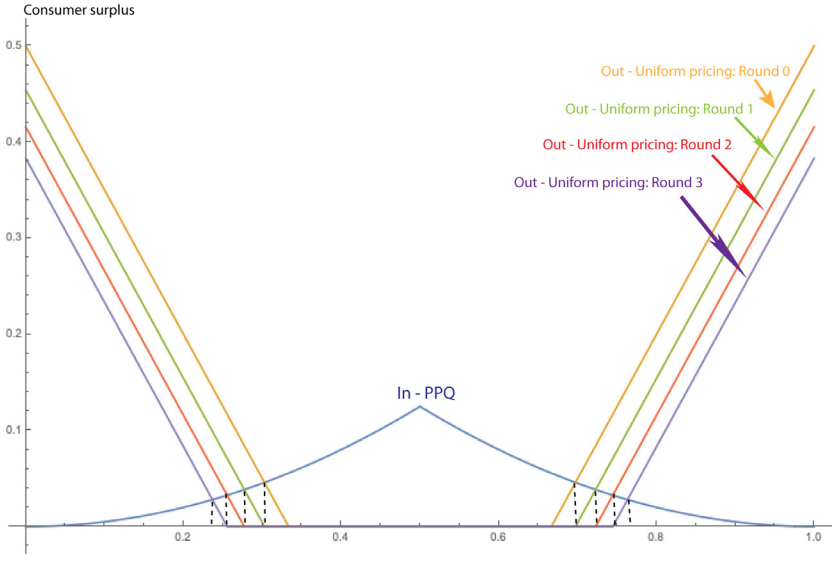


Figure 2 - Consumer surplus in subgame II

3.4 Subgame III: both firms use PPQ for in-consumers, one firm uses NNP for out-consumers, one firm uses uniform pricing for out-consumers

The analysis of this subgame is similar to the previous two: when there exists a in-consumer segment in the middle part of the consumer continuum, firms act as local monopolies in the out-segment and solve the respective optimisation problems as out-lined in the precious analysis. When consumers move first, the consumers close to the middle opt in to receive *PPQ* offers from which they derive higher surplus while loyal consumers opt out to receive *NNP*/uniform pricing offers. Then firms subsequently update their price-quality (schedule) and screening process begins as more consumers opt in to receive *PPQ* offers for the offers in the out-segment being worse than that in in-segment. Following a similar manner, the screening stops until the most loyal consumers are indifferent between the offers in two segments. When

firms move first,to extract maximum rent to maximise profit, given the option to opt out held by consumers, firms set an out-segment offer such that all consumers prefer *PPQ* offers. Hence all consumers opt in to receive *PPQ* offers.

Based on the previous analysis, this argument is immediate. Without unnecessary repetition, we arrive at:

Proposition 3. *When both firms use PPQ for in-consumers, one firm uses NNP for out-consumers, one firm uses uniform pricing for out-consumers, all consumers opt into receive PPQ offer in equilibrium.*

Notice when both firms have personalisation programme i.e. use *PPQ* for in-consumers, same equilibrium will be reached for any other pricing strategy chosen for out-consumer segment.

3.5 Subgame IV: one firm uses PPQ for in-consumers and NNP for out-consumers, the other firm uses NNP for all consumers

In this subgame, we consider the case where only one firm has personalisation pro-gramme. The other firm could be either not having the access to the technology or choosing to forego personalisation. Unlike the subgames where symmetrical pricing strategies were chosen, it is not immediately clear that which consumers are better off with *NNP*/*PPQ* in this case. G09 presented when one firm uses *PPQ* and one firm uses *NNP*, some consumers of the firms using *NNP* is poached by the firm using *PPQ*. This leads to an inefficient equilibrium as a proportion of consumers accept personalised deal from the further firm. Therefore, social welfare is at the lowest in this scenario. In this subsection, we depart from G09 and discuss how market evolves under data-protection regulation.

Consistent to G09 and prior literature, we set the sequence as *PPQ* firm deciding on price-quality schedules after the *NNP*

firm. This is to make the equilibrium analytically tractable as there is no pure strategy Nash equilibrium when pricing is simultaneously set. Hence, *NNP* firm announces its menu of price and quality first and then *PPQ* firm sends perfectly targeted personalised offer to all consumers, who make the purchase decision upon comparing the consumer surplus. In this game, the outside options for *PPQ* firm's consumers include opting out of the personalisation programme and receive *NNP* offer from the *PPQ* firm in addition to purchasing from the *NNP* firm and not purchasing. Without loss of generality, we assume firm L to be the *NNP*-only firm andonly firm R has personalisation programme.

Firm R's optimisation problem is the same as in the both firms using *PPQ* case,formally:

$$(4) \quad \max_{q^R(\theta), s^R(\theta)} \pi^R(\theta), \quad \text{where} \pi^R(\theta) = \theta q^R(\theta) - s^R(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha}, \forall \theta \in [0,1]$$

Firm R sets price, effectively consumer surplus $s^R(\theta)$, to match each consumer's surplus derived from outside option. Firm R offers socially optimal quality due to perfect targeting to maximise profits. Therefore the marginal consumer who feels indifferent between *NNP* offer from firm L and *PPQ* offer from firm R will receive socially optimal surplus from firm R, that is socially optimal quality purchased at marginal cost. Firm L has to offer equal amount of surplus or the consumer is poached otherwise. Hence, the optimisation problem for firm L remains the same as in subgame 1 except IR2 is replaced by the socially optimal surplus curve of firm R:

$$(5) \quad s^L(\theta) = \max_{q^R(\theta)} \left[\theta q^R(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha}\right]\Big|_{\theta=B}$$

Disregarding the opting out option, G09 gives the solution of the optimisation problemsas follows, with marginal consumer type given by $B = [(\frac{2\alpha-1}{\alpha-1})^{\frac{\alpha-1}{\alpha}} + 2]^{-1}$:

$$q^L(\theta) = \begin{cases} (1-2\theta)^{1/(\alpha-1)} & \text{if} \quad \theta \in [0, B] \\ q^l(B) \quad \text{if} \quad \theta \in (B, 1] \end{cases}$$

$$q^R(\theta) = \theta^{1/(\alpha-1)}, \theta \in [0, 1]$$

$$s^L(B) = \left(1 - \frac{1}{\alpha}\right)(1-2B)^{\alpha/(\alpha-1)} - B^{\alpha/(\alpha-1)}$$

$$s^L(\theta) = s^L(B) - \frac{\alpha-1}{2\alpha}(1-2B)^{\alpha/(\alpha-1)} + \frac{\alpha-1}{2\alpha}(1-2\theta)^{\alpha/\alpha-1}, \theta \in [0, B]$$

$$s^R(\theta) = \max[0, (1-\theta)q^L(B) - p^L(B)], \theta \in [B, 1]$$

$$p^L(\theta) = \begin{cases} (1-\theta)(1-2\theta)^{1/(\alpha-1)} - s^L(\theta) & \text{if} \quad \theta \in [0, B] \\ p^L(B) \qquad \text{if} \quad \theta \in (B, 1] \end{cases}$$

$$p^R(\theta) = \begin{cases} 0 \quad \text{if} \quad \theta \in [0, B) \\ \theta^{\alpha/(\alpha-1)} - s^R(\theta) \quad \text{if} \quad \theta \in [B, 1] \end{cases}$$

From this solution, we can obtain that firm R seizes more than half of the market as $B < 1/2$. It can also be shown that more than half of total consumers have their surplus fully extracted by firm R. The only consumers of firm R with positive surplus are those located sufficiently close to B: firm R has to offer certain surplus that is equal to the surplus derived from purchasing the offer for B from firm L to poach those consumers. This is a crucial point when we add in the outside opportunity of opting out of firm R's personalisation programme.

When consumers receive zero surplus, they are indifferent between deal and no-deal. It follows that they should be indifferent between opting in and opting out. If consumers move first, at least for the consumers located at the right of 1/2, opting out of the personalisation programme would prompt firm R to offer a menu of *NNP* price-quality pairs as in subgame 1, from which consumers derive strictly positive surplus. Even if firm moves first and preemptively set an *NNP* menu that no consumer prefers, opting out is still a weakly dominant strategy as they don't receive any surplus

in the in-consumer segment. Firm would have to update price-quality schedules to cover the out-consumer segment to maximise profits. However, the consumers located between 1/2 and B, those who are poached by firm R through *PPQ*, wouldn't benefit from opting out of firm R's personalisation programme. From Lemma 3, we learn that under *NNP*, the optimal quality offered decreases in consumer type and is equal to 0 when $\theta = 1/2$ to suffice IC condition. Therefore, firm R does not provide any *NNP* offers that generate positive surplus to the consumers located in $[0, 1/2]$. Then the outside opportunities of those poached consumers boil down to the *NNP* offers from firm L, more precisely the price-quality designed for marginal consumer B. As per the solution above, firm R sends perfectly targeted offers that exactly match the surplus derived from purchasing $(p^L(B), q^L(B))$. Thus, there is no profitable deviation for those consumers. Note, to opt out and prompt firm L to offer *NNP* menu as in both-*NNP* case is not an option even though all poached consumers would receive strictly higher surplus. This is because firm R could potentially offer socially optimal surplus due to perfect targeting to poach those consumers while firm L has to offer quality schedule that satisfies IC condition. In other words, it's optimal to forego those consumers when the rival has *PPQ*. Hence firm L will never offer a *NNP* menu that covers half of the market.

Simulating with $\alpha = 2$, we can see the above argument in figure 3. Consumers in $[0, 0.27]$ $(B=0.27)$ purchase *NNP* offers from firm L; consumers in $[0.27, 0.5]$ purchase *PPQ* offers from firm R while consumers in [0.5, 1] opt out of the personalisation programme and receive *NNP* offers from firm R. As $\alpha \to 1$, $B \to 1/3$; $\alpha \to \infty$, $B \to 1/4$, provided $\frac{dB}{d\alpha} < 0$, we have $1/4 < B < 1/3$. From $s^R(\theta) = max[0, (1-\theta)\, q^L(B) - p^L(B)]$, we can obtain $\hat{\theta}$ such that $(1-\hat{\theta})q^L(B) - p^L(B) = 0$ and consumers located on the right of $\hat{\theta}$ have surplus fully extracted under *PPQ*. It's straightforward to show that $\hat{\theta} < 1/2$, hence consumers located in $[1/2, 1]$ receive zero surplus under *PPQ* and have incentives to opt out as aforementioned. The entire analysis applies when reversing positions of two firms.

Therefore, we arrive at

Proposition 4. *When one firm uses PPQ for in-consumers and NNP for out-consumers,the other firm uses NNP for all consumers: (1) all consumers located on the half of market on the side of the firm with personalisation programme opt out to receive NNP offers. (2) Part of the consumers located on the other half of the market opt into the personalisation programme of the further firm to receive PPQ offers. (3) The rest of the consumers purchase NNP offers from the closer firm.*

3.6 Subgame V: one firm uses PPQ for in-consumers and NNP for out-consumers, the other firm uses uniform pric-ing for all consumers

The majority of the analysis is analogous to subgame IV. We also assume firm R to be the firm with personalisation programme and sets price-quality schedules after firm L for the same reasoning. From Lemma 4, we have the consumer surplus function when consumers purchase the optimal uniform pricing offer from firm L, formally:

$$s^L(\theta) = (1-\theta)\Big(\frac{\alpha}{2\alpha-1}\Big)^{\frac{1}{1-\alpha}} - \Big(\frac{\alpha}{2\alpha-1}\Big)^{\frac{2-\alpha}{1-\alpha}}$$

Firm R uses perfect targeting to pick up the part of the market not covered (as seen in subgame II):$[\frac{\alpha-1}{2\alpha-1}, 1]$and poach some of the consumers that would purchase from firm L. The firm achieves so by providing the surplus that exactly match that of purchasing the uniform pricing offer. Therefore we can substitute the surplus function into firm R's objective function:

$$\max_{q^R(\theta), s^R(\theta)} \pi^R(\theta), \quad \text{where} \pi^R(\theta) = \theta q^R(\theta) - s^R(\theta) - \frac{(q^R)^\alpha(\theta)}{\alpha}, \forall \theta \in [0, \frac{\alpha-1}{2\alpha-1}]$$

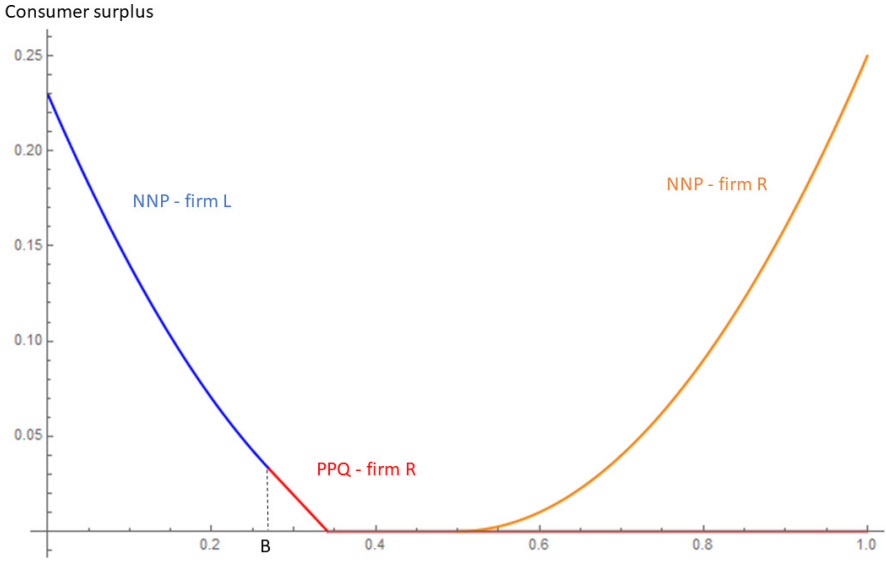


Figure 3 - Consumers surplus in subgame IV

Solving the optimisation problem gives the optimal share of consumers firm R should poach: firm R keeps poaching firm L's consumers until no more profit could be gained. There is no closed form solution for the marginal consumer type B. When α= 2, $B = \frac{-3+\sqrt{13}}{2}$. Then firm L will offer the uniform price and quality such that consumer B is indifferent between purchasing and not purchasing. Therefore, despite different value in B, we arrive at the same layout as in proposition 4:

Proposition 5. *When one firm uses PPQ for in-consumers and NNP for out-consumers,the other firm uses uniform pricing for all consumers: (1) all consumers located on the half of market on the side of the firm with personalisation programme opt out to receive NNP offers. (2) Part of the consumers located on the other half of the market opt into the personalisation programme of the further firm to receive PPQ offers. (3) The rest of the consumers purchase NNP offers from the closer firm.*

4 Discussion

From subgame 1-3, we learn that as long as both firms use *PPQ* for consumers who opt into the personalisation program, the same equilibrium where all consumers opt in will be reached for any pricing stratgies used in out-consumer segment. From subgame4-5, we have that when there is only one firm with personalisation programme, all consumers of that firm will opt out and the firm uses *PPQ* to poach some of the other firm's consumers. Therefore, the choices of pricing strategies for the in-consumer segment determine which equilibrium that the market will reach. Comparing with benchmarked G09, in subgame 1-3, the market equilibrium is the same as that of both firm using *PPQ* case. In subgame 4-5, the equilibrium is different from the one in G09with asymmetrical pricing (one firm uses *NNP* and one firm uses *PPQ*) as half of the total consumers exercise the option of opting out and receive higher surplus instead of being fully extracted. This follows that consumers will only exercise the option to opt out in equilibrium when there is only one firm with personalisation programme. In the cases with asymmetrical pricing strategies (with and without personalisation programme), the total consumer surplus is higher in this model than in G09 despite the social welfare is lower. G09 gives firms' profits to be higher when both firms use *PPQ* comparing to only one firm using *PPQ*. Therefore, it is straightforward to arrive at that firms receive higher profit when they both use *PPQ* in this model as some extra rent under *NNP* will have to be offered to the consumers who opt out when the rival doesn't have personalisation programme.

Intuitively, when both firms have personalisation programme, they are competing for every consumer who opts in hence at least some consumers, especially the less loyal ones, benefit from competition. This, however, is not the case for the more loyal consumers once they opt out of the programme as they no longer receive personalised offers from both firms. Although they are still free to purchase from the price-quality menu offered from both firms, they cannot

derive positive surplus from purchasing *NNP* offer from the further firm as the price-quality is set to prevent product cannibalisation. The two local monopolies are formed in such a way that consumers are better off with the personalised offer. When there is only one firm with personalisation programme, the competition is no longer for every consumer but the consumer at the margin. As firm without *PPQ* couldn't possibly set an *NNP* price quality schedule that attracts the consumers loyal to the rival firm, hence *PPQ* firm could extract all the consumer surplus of its consumers through *PPQ*. Then under data protection regulation, those consumers are better off with opting out of the personalisation programme and claiming information rent under *NNP*.

As aforementioned, when both firms use *PPQ*, consistent with G09, the social welfare is at highest since socially optimal quality is being offered. Data protection regulation does not affect this welfare property. When only one firm has personalisation programme, the option to opt out reduces the social welfare comparing to the scenario with no such option. From propositions 4-5, some of the consumers are poached, which could be seen as an inefficient allocation as it's more efficient for those consumers to purchase from the closer firm. Together with half of the market opting out to receive *NNP* offers with qualities constrained by self-selection, the consumer surplus is higher at the cost of social welfare.

5 Conclusion

Data protection regulation is a relatively new practice while we see the real world applications of the personalisation techniques such as customer relationship management system (CRM) and flexible manufacturing system (FMS) and a clear trend of wider adoption can be seen. Prior literature often flag the strategic responses of the consumers (e.g., deleting cookies; delaying purchases) as a difficult aspect to be incorporated in the models. Regulations such as GDPR offer a firmer instrument for consumers to reject personalisation pricing, which is unprecedented. The novel-

ty of this work consists of providing a theoretical framework to analyse the implications of empowering consumers with such regulation-implied option on the markets that involve-such price-quality personalisation. This effectively embeds the strategic behaviour of the consumers' into the modelling. In the stylised setting, the model shows in some cases, the option to reject personalisation may not increase consumer welfare as the result of a screening and in some cases it might lead to inefficient allocation and a decrease of social welfare. Typically, those need to be assessed when evaluating the merits of the policy. For example, under what circumstances it is worth trading in certain amount of total welfare for the redistribution of surplus; how a mutually beneficial mechanism should be designed to maintain the competition in the in-consumer segments such that optimal social surplus is achieved, etc. At the same time, data regulations generally concern beyond the mere pecuniary consequences on the markets. They are more of a legislation to prevent the abuse of manipulative power of mass data processing, which could serve the advantages of the few. Th is means that maximising economic welfare might not be the sole purpose that legislators are questing. Instead the aim is to seek a balance between the socio-economic factors that manage the risk imposed by big data while not distorting markets excessively. How to achieve this is another big topic per se.

Nonetheless, there are some limitations to this model and caveats for real world applications, which present opportunities for future improvements. Firstly, It is assumed that when a consumer opts into the personalisation programme, his/her location be-comes common knowledge known to both firms. Although this assumption could be robust when the information acquisition is extremely easy such that the costs are negligible, in reality, the personalisation programme is often not universal as it involves a stage of firm investing in collecting data from the consumers. For instance, firms might have loyalty scheme to encourage consumers to fill in the membership form by offering discounts on the current purchase. Hence, it is more plausible that a consumer's

location is only known to one firm and that knowledge is not shared with the other firm. Furthermore, as discussed in section 2, there is a cost of information acquisition and it is not in the firm's best interest to acquire fully accurate information of all consumers. It might be the case that a consumer partially reveals privacy and the personalised offer does not perfectly match his/her willingness to pay, which is not taken into account in this model. To incorporate this into the modelling is certainly away to go in the future. Secondly, common to the majority of Hotelling models, the consumers' preferences and locations are assumed to be ex-ante given and fixed. In reality, the purchase decisions are characterised by a variety of factors other than just brand preference and marginal utility for quality, e.g. regret aversion. It is often to difficult to fully capture all the affecting factors, which gives the rise of random utility hypothesis in econometrics. This suggests consumers' preferences should be varying with time instead of staying fixed and the model does not explain how the preferences are set. Although this is beyond the scope of this model, it is an extension to pursue in related framework. Finally, it is also worth extending the model to investigate the firm's motives in adopting pricing strategies. For example, one might generalise the model by altering the distribution of the consumers and assess different levels of profits. This is crucial in understanding the firms' behaviour especially when looking at the real-world market layout.

Assessing Whether People's Locations Predict Attitudes Towards a US Federal Minimum Wage Increase

Saleel Huprikar
University of Pennsylvania

Abstract

The United States federal minimum wage has not increased since 2009, and the debate regarding whether the federal minimum wage should increase continues to intensify. The field of literature on attitudes towards the federal minimum wage has typically examined the breakdown of such attitudes based only on a few variables such as political affiliation, gender, and race. Using Pew Research Center survey data from 2013-2016 and state-level economic data, this examination investigates whether a person's general location (rural or urban) and the economic conditions of the state that the person resides in can have an influence on the person's perspective towards a federal minimum wage increase, while controlling for the person's political ideology, race, gender, income, education, and generation. This examination's argument is two-fold: 1) People's general locations and the associated state-level economic conditions can have an impact on their perspectives towards a federal minimum wage increase; 2) it is inconclusive whether the neoclassical argument on the minimum wage is losing support among the American public. This research uncovers substantial variation in individual-level minimum wage attitudes based on state-

level economic conditions and general location, suggesting that the public may be more satisfied with state or local-level minimum wage policy solutions.

Keywords

Minimum Wage; Attitudes; Rural; State Minimum Wage; State Unemployment Rate; State Poverty Rate; State Cost of Living; State Median Household Income; State Gini Coefficient

Table of Contents

Background Information

The Fair Labor Standards Act of 1938 (FLSA) first established the US federal minimum wage at $0.25[1]. Since the act's inception, the US Congress has increased the federal minimum wage twenty-two times, though amendments throughout the 20th century have specified different minimum wages for certain types of workers[2]. Thus far in the twenty-first century, Congress has raised the federal minimum wage from $5.85 in 2007, to $6.55 in 2008, and finally to $7.25 in 2009. Although the federal minimum wage has not increased since 2009[3], numerous states and municipalities have increased their respective minimum wages above the federal threshold. This move by regional governments has led minimum wage advocates to call for an increase in the federal minimum wage to levels such as $9, $10.10, $12, or even $15. As a result, the minimum wage debate is far from settled.

Neoclassical economics arguments have long posited that increases in the minimum wage will lead to decreases in employment rates[4]. This postulation relies on the theoretical

---

1 "History of Federal Minimum Wage Rates Under the Fair Labor Standards Act, 1938 - 2009," United States Department of Labor: Wage and Hour Division, accessed on July 3, 2017, https://www.dol.gov/whd/Minimum Wage Approval Rating/chart.htm#fn2.

2 Ibid

3 "State Minimum Wage Rates," Labor Law Center, accessed on July 10, 2017, https://www.laborlawcenter.com/state-minimum-wage-rates/.

4 i. Mikra Krasniqi, "Changing Attitudes towards Minimum Wage Debate: How is The Neoclassical Economic Theory holding in the face of a New Era of Minimum Wage Studies?" George Mason University School of Public Policy, 2007, pg. 3.
ii. Dale Baleman and Paul J. Wolfson, What does the Minimum Wage Do? (Kalamazoo, MI: W.E. Upjohn Institute for

microeconomics competitive labor market model, which shows that a labor market inefficiency occurs with the implementation and increase of a minimum wage (a wage floor). The model demonstrates that raising the minimum wage will increase costs for employers, decrease the demand for workers, and create a surplus of workers who subsequently become unemployed and could end up in poverty.[5] As a result, the neoclassical argument maintains that a minimum wage should not exist, and that wages in the labor market should be determined by the forces of supply and demand in order to ensure that there is no loss of social welfare.[6] However, Mikra Krasniqi (2007) claims that from 1990 to 2006, the majority of scholarly attitudes had started to shift away from the neoclassical argument on minimum wages and move towards the assertion that increases in the minimum wage may not adversely affect employment rates.[7] In fact, more research studies have concluded that modest increases in the minimum wage are not always associated with statistically significant decreases in employment rates.[8] Apart from the employment aspect, minimum wage research has also begun to consider the effects of minimum wage increases on other elements of society, such as income inequality and poverty rates. Early research has shown that higher minimum wages reduce income inequalities and poverty rates.[9]

These challenges to the neoclassical paradigm of the minimum wage have intensified the political debate on the merits of raising the minimum wage (whether it be at the federal, state, or local level). Throughout much of the twentieth century, the politics of the minimum wage

---

Employment Research, 2014), pg. 22.

5 Baleman, What does the Minimum Wage Do?, pg. 11.

6 Ibid, pg. 11.

7 Krasniqi, pg. 2.

8 Baleman, What does the Minimum Wage Do?, pg. 60.

9 Ibid, pgs. 304 and 319

had focused on disagreements between Republicans and Democrats over the magnitude of minimum wage increases and the coverages.[10] The debate over the magnitude of wage change persists in twenty-first century political discourse, but the politics have become more ambiguous because the disagreements are rooted in not only inter-party affairs but also in intra-party affairs. One recent comprehensive minimum wage legislation to be introduced in Congress was the Minimum Wage Fairness Act in 2013 (in the Senate), which proposed raising the US federal minimum wage to $10.10 over the course of two years.[11] Republicans opposed the bill and the Senate did not vote on it, because even though the Democrats controlled the Senate at the time, there were disagreements among Democrats over the magnitude of the proposed federal minimum wage.[12] Advocates for an increase in the federal minimum wage have presented a myriad of arguments; the following ones are some of the most common: 1) more people will be lifted out of poverty; 2) income inequality will lessen; 3) government welfare spending will decrease; and 4) more Americans will be able to afford housing and other basic goods.[13] Opponents of an increased minimum wage tend to argue the following: 1) more workers will be unemployed (especially low-skill ones) due to company layoffs and/or outsourcing; 2) small businesses cannot tolerate a higher minimum wage; 3) poverty rates will increase; and 4) the prices of consumer goods will increase.[14]

10 Jerold Waltman, Politics of the Minimum Wage (Urbana and Chicago: University of Illinois Press, 2000), pgs. 20, 21, 40.

11 "S.1737 - Minimum Wage Fairness Act," Library of Congress, accessed July 17, 2017, https://www.congress.gov/bill/113th-congress/senate-bill/1737.

12 Justin Sink, "Obama: Congress has 'clear choice' on minimum wage," The Hill, April 2, 2014, accessed July 15, 2017, http://thehill.com/blogs/blog-briefing-room/news/202475-obama-congress-has-clear-choice-on-minimum-wage.

13 "Should the Federal Minimum Wage Be Increased?" ProCon, accessed on July 15, 2017, https://minimum-wage.procon.org/.

14 Ibid

The other dimension of the general minimum wage debate concerns whether a minimum wage should exist in the first place. Some who are against the existence of a minimum wage contend that the market should decide the wages of workers (the neoclassical argument), while others argue for minimum wage alternatives such as the institutionalization of universal basic income.

I Literature Review and Project Introduction

The current relevance of the minimum wage debate, the increase in empirical minimum wage research, and changes in minimum wage policy at the state level necessitate an analysis of people's attitudes towards a federal minimum wage increase. Krasniqi (2007) studied scholarly attitudes towards the prospect of a minimum wage increase, but it is also important to examine public opinion towards a minimum wage increase because these attitudes can have an impact on policy. Waltman and Pittman (2002) studied the determinants of state minimum wage rates using survey data, and concluded that the public's political leanings majorly influence state minimum wage levels.[15] Evidence from the 20th century showed that there was a strong connection between US federal minimum wage policy and public opinion, such that Congress tended to increase the federal minimum wage in light of polls that demonstrated that a clear majority of the American public favored a higher federal minimum wage. [16]

Waltman (2000) analyzed public opinion polls in the 1990s and observed that the political affiliation of people is a significant predictor of their support for a federal minimum

15 Jerold Waltman and Sarah Pittman, "The Determinants of State Minimum Wage Rates: A Public Policy Approach," Journal of Labor Research 23, no. 1 (2002), 51-56.

16 Robert S. Erikson and Kent L. Tedin. American Public Opinion: Its Origins, Content and Impact. 9th ed. New York: Routledge, 2015, pgs. 309-310.

wage increase.[17] In addition to political affiliation, numerous polls that have been conducted by major news sources (e.g. CBS News & The Hill) and polling organizations (e.g. Pew Research Center and Gallup Poll) have uncovered that gender and race also influence people's attitudes towards a federal minimum wage increase.[18] However, most of the written content on the poll findings tends to focus on these factors and the overall percentage of people who support a federal minimum wage increase. Much of the written content offers little substance on support levels for a higher federal minimum wage based on the respondent's location and the location's associated economic conditions. There seems to be a dearth of literature that focuses on public opinion towards the federal minimum wage over multiple years since 2009.

The locations of Americans are important to consider because numerous states and local areas have minimum wages that are higher than the federal level of $7.25.[19] In addition to the fact that minimum wages vary from one location to another, economic factors often associated with the minimum wage (e.g. unemployment rate, poverty rate, cost of living, median household income level, and the Gini coefficient[20]) also vary from one location to another. Whether a person lives in a rural or urban area could play a significant role in influencing opinion towards a federal minimum wage increase; rural and urban areas tend to have different economic conditions, such as a difference in the cost of living. Therefore, this paper theorizes that people's locations and the locations' economic conditions may influence people's attitudes towards the prospect of raising the federal minimum wage. This study will only consider state-level economic values of minimum wage, unemployment rate, poverty rate, cost of living, median household income level, and the Gini coefficient. The reason for this is that the survey datasets (which are described in

19 "State Minimum Wage Rates," Labor Law Center, accessed on July 10, 2017, https://www.laborlawcenter.com/state-minimum-wage-rates/.

20 The Gini coefficient is a measure of income inequality that falls between 0 and 1, where 0 means perfect equality and 1 means perfect inequality.

the next section) do not indicate the zip codes of the respondents; the surveys only indicate the states that the respondents reside in. The examination, however, will incorporate people's general locations--rural or urban--since the survey datasets provide this information for each respondent; this general location variable can serve as a proxy for the local-level economic conditions.

This paper seeks to fill the literature void, as described above, by observing whether a person's general location (rural or urban) and the state-level economic conditions of a person's state of residence can have an influence on attitudes towards a federal minimum wage increase, while controlling for political ideology, race, gender, income, education and generation type. The following list of location variables will be scrutinized: a person's general location type (rural or non-rural), state minimum wage, state unemployment rate, state poverty rate, state cost of living, state median household income, and state Gini coefficient. This project examines people's attitudes towards the federal minimum wage from 2013-2016 using survey data. This study will not focus on the debate of whether a federal minimum wage should exist, because the debate surrounding the magnitudes of the federal minimum wage is more persistent. The project has two principal objectives: 1) To determine whether the location-based variables have an impact on one's attitude towards a federal minimum wage increase; 2) To assess whether support for the neoclassical economics argument on the minimum wage, as detailed earlier, is eroding among Americans since the last change of the federal minimum wage in 2009.

With this background information in mind, the examination will provide an overview of the rest of the paper. Section II will describe the Pew Research Center surveys and economic datasets that this research utilizes, their sources, and potential biases. Section III will provide an overview of the main statistical methodologies that this research study employs—ordered and binary logit regressions. Section IV will provide the results of those regressions. Section V will analyze and discuss the important results from Section IV. Section VI will

conclude the examination by addressing the primary objectives from Section I, explaining the applications of the study's findings, and suggesting avenues for future research.

II Data

This project utilizes four Pew Research Center respondent surveys[21]: **February 2013 Political Survey** (n = 1504), **January 2014 Political Survey** (n = 1504), **December 2015 Political Survey** (n = 1500), and **August 2016 Political Survey** (n = 2010). All surveys include the following characteristics of the respondents: gender, race, general type of location, state location, income level, education level, political ideology, and generation type. The February 2013 Political Survey asks whether the federal minimum wage should be raised to $9 (four options for respondents: strongly oppose, oppose, favor, strongly favor). The January 2014 Political Survey asks whether the federal minimum wage should be raised to $10.10 (four options for respondents: strongly oppose, oppose, favor, strongly favor). The December 2015 Political Survey asks whether the federal minimum wage should be increased from $7.25 with no proposed federal minimum wage given (two options for respondents: favor, oppose). The August 2016 Political Survey asks whether the federal minimum wage should be raised to $15 (two options for respondents: favor, oppose).

A dataset was made that contains annual average state-level economic data from 2012-2015. For all fifty states and the District of Columbia, there are measurements for the following economic factors: state minimum wage, state unemployment rate, state poverty rate, state cost of living, state median household income, and the state Gini coefficient. The data on the state minimum wages comes from the Labor Law Center, data on the state cost of living comes from the Missouri Economic Research and Information Center (MERIC)[22], state

21 NOTE: The Pew Research Center bears no responsibility for the interpretations presented or conclusions reached based on analysis of the data in this project.

22 MERIC determines the aggregate cost of living index

unemployment rate data comes from the Bureau of Labor Statistics, and the rest of the data comes from the annual American Community Survey Briefs provided by the US Census Bureau. This economic dataset was merged with each of the four Pew Research Center surveys so that each survey respondent would have economic measures associated with him/her based on his/her state location. This allows for the testing of the hypothesis that the economic conditions of the respondent's state of residence have an impact on the respondent's attitudes towards a federal minimum wage increase.

In terms of biases, response bias will certainly be present in this examination since the wording of each minimum wage question varies from one survey to another. The attitudinal bias that would result from a change in the federal minimum wage in a time period is not present in this examination since the federal minimum wage did not change in the 2013-2016 period.

III Statistical Procedures

III.i Ordered and Binary Logit Frameworks

Two ordered logit and two binary logit regressions are generated in this study. The *MINIMUM WAGE APPROVAL RATING* variable, in each of the four regressions, is the dependent variable, which gives one's response to the minimum wage question of the given survey. Table 1 describes the specifications of the *MINIMUM WAGE APPROVAL RATING* variable in each of the regressions:

for a state based on costs of grocery, housing, utilities, transportation, health, and other miscellaneous items.

Regression	**Regression Type**	**Dependent Variable**	**Domain of Dependent Variable**
2013 ($9)	Ordered logit	MINIMUM WAGE APPROVAL RATING 2013 ($9)	{0,0.333,0.667,1} for strongly oppose, oppose, favor, and strongly favor, respectively
2014 ($10.10)	Ordered logit	MINIMUM WAGE APPROVAL RATING 2014 ($10.10)	{0,0.333,0.667,1} for strongly oppose, oppose, favor, and strongly favor, respectively
2015 (General)	Binary logit	MINIMUM WAGE APPROVAL RATING 2015 (General)	{0,1} for oppose and favor, respectively
2016 ($15)	Binary logit	MINIMUM WAGE APPROVAL RATING 2016 ($15)	{0,1} for oppose and favor, respectively

Table 1

The independent variables of the regressions are defined as follows in Table 2:

Independent Variable	**Description**
State Minimum Wage 201X *	State Minimum Wage in Year 201X (in dollars)
State Unemployment Rate 201X *	State Unemployment Rate in Year 201X (percent)
State Poverty Rate 201X *	State Poverty Rate in Year 201X (percent)
State Cost of Living 201X *	State Cost of Living Index in Year 201X
State Median Household Income 201X *	State Median Household Income in Year 201X (in thousand-dollars)
State Gini Coefficient 201X *	State Gini Coefficient in Year 201X (percent)
Rural *	Binary (1 for Rural, 0 for Non-Rural)
Female	Binary (1 for Female, 0 for Male)
White	Binary (1 for White, 0 for Non-White)
Black	Binary (1 for Black, 0 for Non-Black)
Hispanic	Binary (1 for Hispanic, 0 for Non-Hispanic)
Income	Income level of respondent; scale: 1-9 1: less than $10K 2: $10K-$19K 3: $20K-$29K 4: $30K-$39K 5: $40K-$49K 6: $50K-$74K 7: $75K-$99K 8: $100K-$149K 9: $150K and higher
Education	Education level of respondent; scale: 1-8 1: Less than High School 2: High School incomplete 3: High School graduate 4: Some college and no degree 5: Two-year associate degree 6: 4-year college or univer- sity degree/Bachelor's degree 7: Some postgraduate or professional schooling and no degree 8: Postgraduate or professional degree
Political Ideology	Political Ideology of respondent; scale: 1-5 1: Very conservative 2: Conservative 3: Moderate 4: Liberal 5: Very liberal
Generation	Respondent's Generation; scale: 1-5 1: Millennials (born in 1981-) 2: Xer (born in 1965-1980) 3: Baby Boomers (born in 1946-1964) 4: Silent (born in 1928-1945) 5: Greatest and older (born in -1927))

Table 2

* denotes a location variable

Note: Tables 4-7 in the appendix show the weighted summary statistics for the location variables in each of the four surveys.

III.ii Independent and Dependent Variables

All four regressions[23] are weighted in order to account for potential biases in sampling and to better reflect certain US population characteristics. Each Pew Research Center dataset provides a survey weight variable that will be utilized for each regression.

The 2013 and 2014 ordered logit models have the following forms (for each model, 3 equations for each value j):

$$\ln\left(\frac{Prob(MINIMUM\ WAGE\ APPROVAL\ RATING\ 2013 \leq j)}{1-Prob(MINIMUM\ WAGE\ APPROVAL\ RATING\ 2013 \leq j)}\right) = \alpha_j + (-\beta_1(State\ Minimum\ Wage\ 2013) - \beta_2(State\ Unemployment\ Rate\ 2012) - \beta_3(State\ Poverty\ Rate\ 2012) - \beta_4(State\ Cost\ of\ Living\ 2012) - \beta_5(State\ Median\ Household\ Income\ 2012) - \beta_6(State\ Gini\ Coefficient\ 2012) - \beta_7(Rural) - \beta_8(Female) - \beta_9(White) - \beta_{10}(Black) - \beta_{11}(Hispanic) - \beta_{12}(Income) - \beta_{13}(Education) - \beta_{14}(Political\ Ideology) - \beta_{15}(Generation)) + \varepsilon$$

23 For each model, the state unemployment rate, state poverty rate, state cost of living, state median household income, and the state Gini coefficient will take on their respective annual average values in the year before the administration of the corresponding Pew survey in order to prevent post-treatment bias. The state minimum wage will be the minimum wage in the state at the time the corresponding Pew survey had been administered. None of the models exhibit multicollinearity. Neither the 2013 ordered logit regression nor the 2014 ordered logit regression shows a violation of the Proportional Odds assumption that would adversely affect model specification. Neither the 2015 binary logit regression nor the 2016 binary logit regression provides evidence that the continuous independent variables are not linearly related to the log odds dependent variable. As a result, it is appropriate to proceed with these models for the rest of the study.

$$\ln\left(\frac{Prob(MINIMUM\ WAGE\ APPROVAL\ RATING\ 2014 \leq j)}{1-Prob(MINIMUM\ WAGE\ APPROVAL\ RATING\ 2014 \leq j)}\right) = \alpha_j +$$

$$(-\beta_1(\text{State Minimum Wage 2014}) - \beta_2(\text{State Unemployment Rate 2013}) - \beta_3(\text{State Poverty Rate 2013}) - \beta_4(\text{State Cost of Living 2013}) - \beta_5(\text{State Median Household Income 2013}) - \beta_6(\text{State Gini Coefficient 2013}) - \beta_7(\text{Rural}) - \beta_8(\text{Female}) - \beta_9(\text{White}) - \beta_{10}(\text{Black}) - \beta_{11}(\text{Hispanic}) - \beta_{12}(\text{Income}) - \beta_{13}\text{Education}) - \beta_{14}(\text{Political Ideology}) - \beta_{15}(\text{Generation})) + \varepsilon$$

j {0,0.333,0.667} and j represents the cut-point (intercept) at category j.

The 2015 and 2016 binary logit models have the following forms:

$$\ln\left(\frac{Prob(MINIMUM\ WAGE\ APPROVAL\ RATING\ 2015 = 1)}{1-Prob(MINIMUM\ WAGE\ APPROVAL\ RATING\ 2015 = 1)}\right) = \beta_0 +$$

$$\beta_1(\text{State Minimum Wage 2015}) + \beta_2(\text{State Unemployment Rate 2014}) + \beta_3(\text{State Poverty Rate 2014}) + \beta_4(\text{State Cost of Living 2014}) + \beta_5(\text{State Median Household Income 2014}) + \beta_6(\text{State Gini Coefficient 2014}) + \beta_7(\text{Rural}) + \beta_8(\text{Female}) + \beta_9(\text{White}) + \beta_{10}(\text{Black}) + \beta_{11}(\text{Hispanic}) + \beta_{12}(\text{Income}) + \beta_{13}(\text{Education}) + \beta_{14}(\text{Political Ideology}) + \beta_{15}(\text{Generation}) + \varepsilon$$

$$\ln\left(\frac{Prob(MINIMUM\ WAGE\ APPROVAL\ RATING\ 2016 = 1)}{1-Prob(MINIMUM\ WAGE\ APPROVAL\ RATING\ 2016 = 1)}\right) = \beta_0 +$$

$$\beta_1(\text{State Minimum Wage 2016}) + \beta_2(\text{State Unemployment Rate 2015}) + \beta_3(\text{State Poverty Rate 2015}) + \beta_4(\text{State Cost of Living 2015}) + \beta_5(\text{State Median Household Income 2015}) + \beta_6(\text{State Gini Coefficient 2015}) + \beta_7(\text{Rural}) + \beta_8(\text{Female}) + \beta_9(\text{White}) + \beta_{10}(\text{Black}) + \beta_{11}(\text{Hispanic}) + \beta_{12}(\text{Income}) + \beta_{13}(\text{Education}) + \beta_{14}(\text{Political Ideology}) + \beta_{15}(\text{Generation}) + \varepsilon$$

IV Results

This table reports the results of the regressions as outlined in Section III:

	MINIMUM WAGE APPROVAL RATING 2013 ($9) Ordered logit	MINIMUM WAGE APPROVAL RATING 2014 ($10.10) Ordered logit	MINIMUM WAGE APPROVAL RATING 2015 (General) Binary logit	MINIMUM WAGE APPROVAL RATING 2016 ($15) Binary logit
State Minimum Wage	0.054 (0.075)	0.041 (0.069)	0.058 (0.085)	0.066 (0.057)
State Unemployment Rate	-0.016 (0.031)	-0.022 (0.033)	-0.012 (0.051)	0.128** (0.056)
State Poverty Rate	0.104*** (0.027)	-0.136*** (0.026)	-0.026 (0.038)	-0.102*** (0.036)
State Cost of Living	-0.010** (0.004)	0.008** (0.003)	0.010* (0.006)	0.002 (0.004)
State Median Household Income	0.044*** (0.010)	-0.041*** (0.009)	-0.008 (0.012)	0.008 (0.011)
State Gini Coefficient	-0.053** (0.025)	0.111*** (0.025)	0.039 (0.035)	0.092*** (0.028)
Rural	-0.320*** (0.085)	0.035 (0.080)	0.041 (0.105)	-0.264*** (0.090)
Female	0.410*** (0.063)	0.424*** (0.059)	0.348*** (0.076)	0.458*** (0.067)
White	-0.542*** (0.129)	-0.746*** (0.120)	-0.373** (0.145)	-0.421*** (0.118)
Black	0.892*** (0.161)	0.081 (0.147)	1.059*** (0.216)	1.478*** (0.170)

Table 3

Hispanic	-0.338** (0.150)	-0.256* (0.139)	0.245 (0.180)	1.334*** (0.152)
Income	-0.031** (0.015)	-0.105*** (0.014)	-0.102*** (0.018)	-0.191*** (0.016)
Education	-0.091*** (0.021)	0.006 (0.018)	-0.116*** (0.023)	0.053** (0.021)
Political Ideology	0.655*** (0.036)	0.585*** (0.033)	0.691*** (0.040)	0.752*** (0.036)
Generation	0.010 (0.032)	0.030 (0.030)	0.147*** (0.039)	0.301*** (0.035)
Intercept	-	-	-2.572* (1.326)	-6.648*** (1.128)
0|0.333 cutpoint (Intercept)	-0.477 (1.100)	0.315 (1.068)	-	-
0.333|0.667 cutpoint (Intercept)	0.722 (1.101)	1.608 (1.068)	-	-
0.667|1 cutpoint (Intercept)	2.542** (1.101)	3.278*** (1.069)	-	-
Number of Observations	1216	1259	1259	1740
McFadden R^2	0.2199	0.2053	0.2443	0.3124

Table 3 Cont.

Notes:
* Statistically significant at the 10 percent level
** Statistically significant at the 5 percent level
*** Statistically significant at the 1 percent level

V Analysis

V.i State-Level Variables Discussion

The four regressions show that state-level economic conditions are relevant and may influence a respondent's position on whether the federal minimum wage should increase. Compared to the 2015 binary logit regression (where the survey question asks whether the federal minimum wage should increase at all), the other three regressions show more statistically significant results for the location variables, perhaps because the survey questions for the other three regressions provide proposed federal minimum wages ($9, $10.10, and $15). The state minimum wage seems to have no bearing on one's attitude towards a federal minimum wage increase. This could be attributed to the fact that a state minimum wage usually does not have much economic impact, and is instead more symbolic in meaning.[24] Based on this, the state minimum wage and its associated effects may not be salient to people when they consider their positions on the prospect of a federal minimum wage increase.

The regressions for $10.10 and $15 provide evidence of a negative relationship between the state poverty rate and the support for an increased federal minimum wage, but the regression for $9 shows the opposite. The data shows that for a 1 percent increase in the state poverty rate, the odds of opposing a $9 federal minimum wage decrease by 9.87 percent[25], the odds of opposing a $10.10 federal minimum wage increase

24 Waltman, "The Determinants of State Minimum Wage Rates: A Public Policy Approach," pg. 54.

25 Note that for ordered logit regressions of 2013 and 2014, "odds of opposing" refers to the odds of "strongly opposing" or "opposing" a $9 or $10.10 federal minimum wage relative to the upper two categories — "favor" and "strongly favor".

by 14.54 percent and the odds of opposing a $15 federal minimum wage increase by 9.73 percent. This finding may suggest that people who live in states with high poverty rates believe that an increase in the federal minimum wage can counteract high poverty to a certain extent (around $9) as Figure 1 in the appendix suggests. It is possible that people in high poverty states are more wary of a federal minimum wage that is higher than $9 as Figures 2 and 3 indicate, a point at which they may be more in line with the neoclassical economics argument that an increase in the federal minimum wage will not alleviate poverty.[26] The findings for $10.10 and $15 are intriguing because they contradict a long range of economic studies that argue raising the minimum wage significantly reduces poverty rates; in fact, Dube (2017), using data from 1984-2013, showed that for every 10 percent increase in the effective minimum wage, the poverty rate is expected to decline by 5.3 percent in the long run.[27] There may be a disconnect between the public's perceptions and empirical economic studies on the relationship between a federal minimum wage increase and the state poverty rate.

The 2014 ($10.10) and 2015 (General Support) regressions show a positive relationship between the state cost of living and the support for an increased federal minimum wage, but the regression for $9 shows the opposite. The data shows that for a unit increase in the state cost of living, the odds of op-

26 Thomas C. Leonard, "The Very Idea of Applying Economics: The Modern Minimum-Wage Controversy and its Antecedents," History of Political Economy 32, no. 5 (2000), pg. 124.

27 David Cooper, "Raising the minimum wage to $15 by 2024 would lift wages for 41 million American workers," Economic Policy Institute, April 26, 2017, accessed March 3, 2018. https://www.epi.org/publication/15-by-2024-would-lift-wages-for-41-million/
Arindrajit Dube, "Minimum Wages and the Distribution of Family Incomes," No 10572, IZA Discussion Papers, Institute for the Study of Labor (IZA), 2017.

posing a $9 federal minimum wage increase by 0.95 percent, the odds of opposing a $10.10 federal minimum wage decrease by 0.79 percent, and the odds of opposing an increase in the federal minimum wage (in general) decrease by 1.03 percent. People who live in states with high costs of living generally need higher incomes in order to afford basic necessities (e.g. housing and food) and to have adequate purchasing power. It is intuitive that there is a positive relationship between the support for a higher federal minimum wage in general and the state cost of living. Interestingly, there is a negative relationship between the state cost of living and the support for a $9 federal minimum wage. It could be that those in states with high costs of living believe a $9 federal minimum wage, despite being higher than the current federal minimum wage of $7.25, is not high enough as a living wage, and thus they favor a federal minimum wage higher than $9. On the other hand, those who live in low-cost states may believe that a federal minimum wage of $9 is tolerable.

However, they may also believe that having a federal minimum wage greater than some threshold (perhaps $9) overcompensates for the low cost of living with the potential to harm their economies (e.g. job loss), as suggested in research conducted by the Heritage Foundation.[28]

The regressions for $10.10 and $15 provide evidence of a positive relationship between the state Gini coefficient and the support for an increased federal minimum wage, but the regression for $9 shows the opposite. The data shows that for a 1 percent increase in the state Gini coefficient, the odds of opposing a $9 federal minimum wage increase by 5.41 percent, the odds of opposing a $10.10 federal minimum wage decrease by 10.46 percent, and the odds of opposing a $15 federal minimum wage decrease by 9.67 percent. This finding could sug-

28 James Sherk, "How $15-per-Hour Minimum Starting Wages Would Affect Each State," The Heritage Foundation, August 17, 2016, accessed on December 20, 2018, https://www.heritage.org/budget-and-spending/report/how-15-hour-minimum-starting-wages-would-affect-each-state

gest that people in high income inequality states believe that a $10.10 or $15 federal minimum wage is a sound policy that attempts to alleviate income inequality, while thinking that a $9 federal minimum wage is not sufficient to combat income inequality (Figure 4 captures the relationship between state Gini coefficient and the support for a $9 federal minimum wage). In contrast, those who live in states with low income inequality may view a $10.10 or $15 federal minimum wage as an unsound policy as seen in Figures 5 and 6. People who live in low income inequality states may subscribe to the belief that an increase in the federal minimum wage can reduce income inequality to a certain extent, which is what Litwin (2015) concluded in his study of OECD countries.[29]

The state unemployment rate does not have predictive power in the 2013, 2014 or 2015 regressions. However, it does contain predictive power in the 2016 regression, where for a 1 percent increase in the state unemployment rate, the odds of supporting a $15 federal minimum wage increase by 13.66 percent as Figure 7 depicts. From an employment standpoint, the neoclassical economics argument on the minimum wage seems to be rejected, since there is no perceived negative relationship between the odds of supporting a $15 federal minimum wage and the state unemployment rate. But the state poverty rate, as seen in Figure 3, has a negative relationship with support for a $15 federal minimum wage, leaving the state of the neoclassical economics argument in the realm of public opinion in limbo.

V.ii Rural Discussion

Whether a person lives in a rural area or not is relevant for the 2013 and 2016 regressions. In both cases, the likelihood of supporting a higher federal minimum wage (whether it is $9 or $15) decreases when shifting from non-rural areas to rural

29 Benjamin S. Litwin, "Determining the Effect of the Minimum Wage on Income Inequality," Student Publications (2015), https://cupola.gettysburg.edu/student_scholarship/300

areas. The 2013 ordered logit regression indicates that all else equal, the odds of opposing a $9 federal minimum wage increase by 37.69 percent when one shifts from non-rural to rural. The 2016 binary logit regression indicates that, all else equal, the odds of opposing a $15 federal minimum wage increase by 23.19 percent when one shifts from a non-rural person to a rural person. The rural variable does not have statistical power in the $10.10 case or in the general case, suggesting that the rural variable is sensitive to the extreme ends of the spectrum of proposed federal minimum wages put forth by minimum wage advocates. Economic research has suggested that raising the federal minimum wage has the potential of inflicting harm on rural communities. Rural workers, who tend to receive lower wages than their urban counterparts, are more likely to be affected by dis-employment forces that would result from increases in costs for rural employers.[30] This could explain why rural people, compared to their urban counterparts, are more likely to be against a minimum wage increase.

VI Conclusion

The overarching purpose of this paper is to consider people's locations and the economic conditions associated with those locations when assessing their attitudes towards a federal minimum wage increase. To reiterate, the goals of this project are: 1) To determine whether location-based variables have impacts on a person's attitude towards a federal minimum wage increase; 2) To assess whether the support for the neoclassical economics argument on the minimum wage is eroding among Americans in the post-2009 era.

30 Lisa Marshall, "Minimum-wage hikes could push low-pay workers away," University of Colorado Boulder, June 15, 2017, accessed June 15, 2018, https://www.colorado.edu/today/2017/06/15/minimum-wage-hikes-could-push-low-pay-workers-away

With regards to the first objective, the analysis in Section V provides evidence for the hypothesis that state-level economic conditions and people's general location type (rural or urban) have a bearing on people's viewpoints on the federal minimum wage. The specific location variables that have predictive power in determining one's perspective towards a federal minimum wage increase seem to depend on the proposed federal minimum wage itself (as given in the survey question). For example, in the 2015 regression (which asked whether the federal minimum wage should increase at all), the state cost of living was the only location variable that had predictive power; however, in the 2016 regression (which asked whether the federal minimum wage should be increased to $15), the following location variables had predictive power: state unemployment rate, state poverty rate, state Gini coefficient, and whether the person lives in a rural area. Furthermore, the proposed federal minimum wage can dictate the direction of support for a federal minimum wage increase; with respect to the state poverty rate, the state Gini coefficient, and the state cost of living, the attitudes towards a $9 federal minimum wage switched for $10.10 and $15 federal minimum wages.

With regards to the second objective, results are inconclusive. It cannot be determined if the neoclassical argument on the minimum wage is losing support among the American public. The neoclassical argument on the minimum wage is tied to both unemployment and poverty. In the 2013 and 2014 regressions, the state poverty rate had predictive power whereas the state unemployment rate did not, and the state poverty rate exhibited opposite relationships with the likelihood of supporting a federal minimum wage increase (positive in 2013, negative in 2014). In the 2015 regression, neither the state poverty rate nor the state unemployment rate had statistical power in determining attitudes towards a federal minimum wage increase in general. Then, in the 2016 regression, the state unemployment rate, rather surprisingly, had a positive relationship with the support for a $15 federal minimum wage whereas the state poverty rate had a negative relationship. If the neoclassical economics argument on the minimum

wage had been losing support among members of the general public, there would have been a consistent pattern of positive relationships between the state unemployment/poverty rate and the support for a higher federal minimum wage. Since there was no consistent pattern of positive relationships, it is hard to deduce from the results whether the neoclassical argument is gaining or losing support among Americans in the post-2009 era.

The findings of this study, however, can still help politicians distinguish between constituents who are more likely to support a federal minimum wage increase and those who are less likely to do so. When determining which constituents are more likely to favor an increase in the federal minimum wage (or less likely to do so), politicians need to look beyond political affiliation, gender, and race, and consider the constituents' general location types and the state-level economic conditions of the constituents. Considering that a federal minimum wage policy signifies enforcing a fixed minimum wage across the country (regardless of regional economic differences), the results of this study indicate that the public may not look favorably upon the prospect of formulating minimum wage policy at the federal level. Rather, the public would prefer minimum wage policies to be formulated at a more regional level, either the state level or the local level. This means regional differences in certain economic variables (e.g. cost of living and the level of income inequality) would dictate different minimum wage policies that are tailored to the respective needs of constituents in different regional areas.

Future research in the area of public opinion on the federal minimum wage could examine how the local-level (municipality) economic conditions affect attitudes towards a federal minimum wage or a state minimum wage increase. Also, researchers could assess the American public's level of support for the neoclassical argument on the minimum wage since 2016. Although this study could not determine the state of the neoclassical argument for the American public between 2013-2016, it is possible that researchers may be able to determine

the state of this argument post-2016 due to recent political changes at the federal government level. Researchers could further explore if there is any evidence of dissonance between the American public's opinion on minimum wage and findings from minimum wage empirical research. As seen in the subsection State-Level Variables Discussion in Section V, on the matter of the state poverty rate, there is evidence signalling a disconnect between the American public's opinions on the federal minimum wage and empirical research findings.

See last endnote for the works cited list.[31]

31 2014 American Community Survey 1-Year Estimates, U.S. Census Bureau. Accessed on May 10, 2018. https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk

August 2016 Political Survey. Pew Research Center & American Life Project. Conducted on August 9-16. 2016. Accessed on July 25, 2017.

Baleman, Dale and Paul J. Wolfson. What does the Minimum Wage Do? Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 2014.

Bishaw, Alemayehu and Brian Glassman. "Poverty: 2014 and 2015." U.S. Census Bureau. September 2016. Accessed on August 10, 2017. https://www.census.gov/content/dam/Census/library/publications/2016/demo/acsbr15-01.pdf

Bishaw, Alemayehu and Craig Benson. "Poverty: 2015 and 2016." U.S. Census Bureau. September 2017. Accessed on November 1, 2017. https://www.census.gov/content/dam/Census/library/publications/2017/acs/acsbr16-01.pdf

Bishaw, Alemayehu and Kayla Fontenot. "Poverty: 2012 and 2013." U.S. Census Bureau. September 2014. Accessed on August 10, 2017. https://www.census.gov/content/dam/Census/library/publications/2014/acs/acsbr13-01.pdf

Cooper, David. "Raising the minimum wage to $15 by 2024 would lift wages for 41 million American workers." Economic Policy Institute. April 26, 2017. Accessed March 3, 2018. https://www.epi.org/publication/15-by-2024-would-lift-wages-for-41-million/

"Cost of Living Data Series: 2012 Annual Average." Missouri Economic Research and Information Center (MERIC). Accessed on May 15, 2018.

"Cost of Living Data Series: 2013 Annual Average." Missouri Economic Research and Information Center (MERIC). Accessed on May 15, 2018.

"Cost of Living Data Series: 2014 Annual Average." Missouri Economic Research and Information Center (MERIC). Accessed on May 15, 2018.

"Cost of Living Data Series: 2015 Annual Average." Missouri Economic Research and Information Center (MERIC). Accessed on May 15, 2018.

"Cost of Living Data Series: 2016 Annual Average." Missouri Economic Research and Information Center (MERIC). Accessed on June 10, 2018.

December 2015 Political Survey. Pew Research Center & American Life Project. Conducted on December 8-13, 2015. Accessed on July 25, 2017.

Drake, Bruce. "Polls show strong support for minimum wage hike." Pew Research Center. March 4, 2014. Accessed on January 31, 2018. http://www.pewresearch.org/fact-tank/2014/03/04/polls-show-strong-support-for-minimum-wage-hike/

Dube, Arindrajit. "Minimum Wages and the Distribution of Family Incomes." No 10572. IZA Discussion Papers. Institute for the Study of Labor (IZA). 2017.

Dugan, Andrew. "Most Americans for Raising Minimum Wage. " Gallup Poll. November 11, 2013. Accessed on January 20, 2018. https://news.gallup.com/poll/165794/americans-raising-minimum-wage.aspx

Elis, Niv. "Poll: Bipartisan majority supports raising minimum wage." The Hill. June 1, 2017. Accessed on January 10, 2018. http://thehill.com/homenews/335837-poll-bipartisan-majority-supports-raising-minimum-wage

Erikson, Robert S. and Kent L. Tedin. American Public Opinion: Its Origins, Content and Impact. 9th ed. New York: Routledge, 2015.

February 2013 Political Survey. Pew Research Center & American Life Project. Conducted on February 13-18, 2013. Accessed on July 20, 2017.

Guzman, Gloria G. "Household Income: 2016." U.S. Census Bureau. September 2017. Accessed on November 1, 2017. https://www.census.gov/content/dam/Census/library/publications/2017/acs/acsbr16-02.pdf

"History of Federal Minimum Wage Rates Under the Fair Labor Standards Act, 1938 - 2009." United States Department of Labor: Wage and Hour Division. Accessed on July 3, 2017. https://www.dol.gov/whd/Minimum Wage Approval Rating/chart.htm#fn2

January 2014 Political Survey. Pew Research Center & American Life Project. Conducted on January 15-19, 2014. Accessed on July 20, 2017.

Krasniqi, Mikra. "Changing Attitudes towards Minimum Wage Debate: How is The Neoclassical Economic Theory holding in the face of a New Era of Minimum Wage Studies?" George Mason University School of Public Policy, 2007.

Leonard, Thomas C. "The Very Idea of Applying Economics: The Modern Minimum-Wage Controversy and its Anteced-

ents." History of Political Economy 32, no. 5 (2000): 117-144.

Litwin, Benjamin S. "Determining the Effect of the Minimum Wage on Income Inequality." Student Publications (2015). https://cupola.gettysburg.edu/student_scholarship/300

Marshall, Lisa. "Minimum-wage hikes could push low-pay workers away." University of Colorado Boulder. June 15, 2017. Accessed June 15, 2018. https://www.colorado.edu/today/2017/06/15/minimum-wage-hikes-could-push-low-pay-workers-away

Noss, Amanda. "Household Income: 2012." U.S. Census Bureau. September 2013. Accessed on August 10, 2017. https://www2.census.gov/library/publications/2013/acs/acsbr12-02.pdf

Noss, Amanda. "Household Income: 2013." U.S. Census Bureau. September 2014. Accessed on August 10, 2017. https://www.census.gov/content/dam/Census/library/publications/2014/acs/acsbr13-02.pdf

Posey, Kirby G. "Household Income: 2015." U.S. Census Bureau. September 2016. Accessed on August 10, 2017. https://www.census.gov/content/dam/Census/library/publications/2016/demo/acsbr15-02.pdf

"S.1737 - Minimum Wage Fairness Act." Library of Congress. Accessed July 17, 2017. https://www.congress.gov/bill/113th-congress/senate-bill/1737

Saad, Lydia. "In U.S., 71% Back Raising Minimum Wage." Gallup Poll. March 6, 2013. Accessed on February 20, 2018. https://news.gallup.com/poll/160913/back-raising-minimum-wage.aspx

Sherk, James. "How $15-per-Hour Minimum Starting Wages Would Affect Each State." The Heritage Foundation. August 17, 2016. Accessed on December 20, 2018. https://www.heritage.org/budget-and-spending/report/how-15-hour-minimum-starting-wages-would-affect-each-state

"Should the Federal Minimum Wage Be Increased?" ProCon. Accessed on July 15, 2017. https://minimum-wage.procon.org/

Sink, Justin. "Obama: Congress has 'clear choice' on minimum wage." The Hill. April 2, 2014. Accessed July 15, 2017. http://thehill.com/blogs/blog-briefing-room/news/202475-obama-congress-has-clear-choice-on-minimum-wage

"State Minimum Wage Rates." Labor Law Center. Accessed on July 10, 2017. https://www.laborlawcenter.com/state-minimum-wage-rates/

Tabachnick, B. G. and L. S Fidell. Using multivariate statistics. 6th ed. Boston, MA: Pearson, 2012.
"Unemployment Rates for States, 2012 Annual Averages." Bureau of Labor Statistics. February 27, 2018. Accessed on March 2, 2018. https://www.bls.gov/lau/lastrk12.htm

"Unemployment Rates for States, 2013 Annual Averages." Bureau of Labor Statistics. February 27, 2018. Accessed on March 2, 2018. https://www.bls.gov/lau/lastrk13.htm

"Unemployment Rates for States, 2014 Annual Averages." Bureau of Labor Statistics. February 27, 2018. Accessed on March 2, 2018. https://www.bls.gov/lau/lastrk14.htm

"Unemployment Rates for States, 2015 Annual Averages." Bureau of Labor Statistics. February 27, 2018. Accessed on March 2, 2018. https://www.bls.gov/lau/lastrk15.htm

"Unemployment Rates for States, 2016 Annual Averages." Bureau of Labor Statistics. February 27, 2018. Accessed on March 2, 2018. https://www.bls.gov/lau/lastrk16.htm

Waltman, Jerold. Politics of the Minimum Wage. Urbana and Chicago: University of Illinois Press, 2000.

Waltman, Jerold and Sarah Pittman. "The Determinants of State Minimum Wage Rates: A Public Policy Approach." Journal of Labor Research 23, no. 1 (2002), 51-56.

VII Appendix

VII.i Table 4: Weighted Summary Statistics (2013 Political Survey)

	Minimum	Mean	Standard Deviation	Median	Maximum
MINIMUM WAGE APPROVAL RATING 2013 ($9)	0	0.67	0.33	0.67	1
State Minimum Wage 2013	6.15	7.54	0.48	7.25	9.19
State Unemployment Rate 2012 (percent)	3.1	8	1.47	8.2	11.2
State Poverty Rate 2012 (percent)	10	15.83	2.75	16.4	24.2
State Cost of Living 2012	90.5	104.92	15.02	98.8	167.1
State Median Household Income 2012 (thousands)	37.09	52.21	7.81	50.74	71.12
State Gini Coefficient 2012 (percent)	41.7	46.85	1.80	46.9	53.4
Rural	0	0.17	0.38	0	1

Table 4: Weighted Summary Statistics (2013 Political Survey)

N = 1504

VII.ii Table 5: Weighted Summary Statistics (2014 Political Survey)

	Minimum	**Mean**	**Standard Deviation**	**Median**	**Maximum**
MINIMUM WAGE APPROVAL RATING 2014 ($10.10)	0	0.69	0.32	0.67	1
State Minimum Wage 2014	7.25	7.69	0.50	7.50	9.32
State Unemployment Rate 2013 (percent)	2.9	7.35	1.21	7.6	9.6
State Poverty Rate 2013 (percent)	8.7	15.69	2.69	16	24
State Cost of Living 2013	89.1	105.33	16.33	96.5	156.9
State Median Household Income 2013 (thousands)	37.96	53.64	7.88	51.7	72.48
State Gini Coefficient 2013 (percent)	40.8	47.37	1.84	47.7	53.2
Rural	0	0.17	0.37	0	1

Table 5: Weighted Summary Statistics (2014 Political Survey)

N = 1504

VII.iii Table 6: Weighted Summary Statistics (2015 Political Survey)

	Minimum	Mean	Standard Deviation	Median	Maximum
MINIMUM WAGE APPROVAL RATING 2015 (General)	0	0.75	0.43	1	1
State Minimum Wage 2015	7.25	7.99	0.73	8.05	10.5
State Unemployment Rate 2014 (percent)	2.7	6.15	1.0	6.3	7.9
State Poverty Rate 2014 (percent)	9.2	15.39	2.59	15.9	21.5
State Cost of Living 2014	86.9	105.15	15.77	97.5	164
State Median Household Income 2014 (thousands)	39.68	54.76	8.30	53.03	73.97
State Gini Coefficient 2014 (percent)	41.8	47.27	1.82	47.6	52.2
Rural	0	0.15	0.36	0	1

Table 6: Weighted Summary Statistics (2015 Political Survey)

N = 1500

VII.iv Table 7: Weighted Summary Statistics (2016 Political Survey)

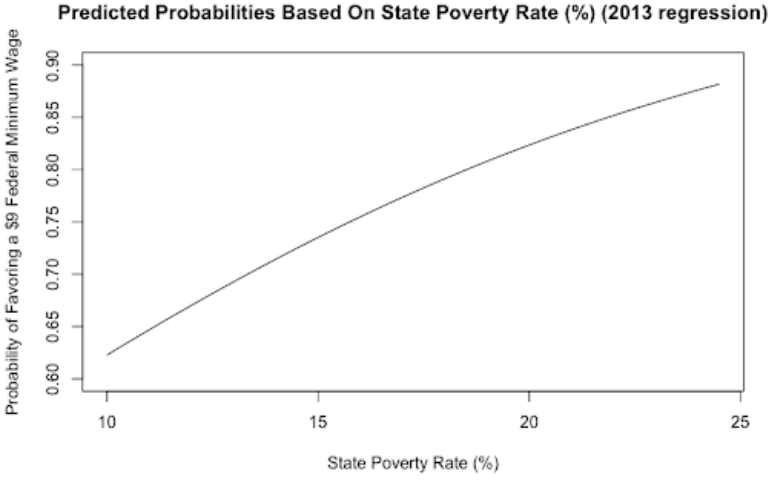
	Minimum	Mean	Standard Deviation	Median	Maximum
MINIMUM WAGE APPROVAL RATING 2016 ($15)	0	0.59	0.49	1	1
State Minimum Wage 2016	7.25	8.19	0.99	8.05	11.5
State Unemployment Rate 2015 (percent)	2.8	5.26	0.81	5.3	6.9
State Poverty Rate 2015 (percent)	8.2	14.7	2.46	15.3	22
State Cost of Living 2015	83.5	106.03	17.89	98.1	168.6
State Median Household Income 2015 (thousands)	40.59	56.63	8.12	55.65	75.85
State Gini Coefficient 2015 (percent)	42.5	47.5	1.84	47.8	53.5
Rural	0	0.18	0.38	0	1

Table 7: Weighted Summary Statistics (2016 Political Survey)

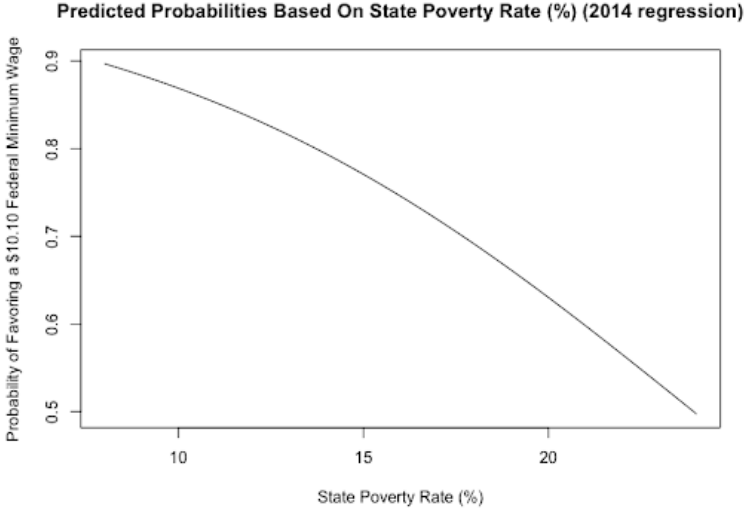
N = 2010

Figures 1-7 on the next page are predicted probability graphs based on one of the location variables. In each graph, all other variables are fixed at their respective means.

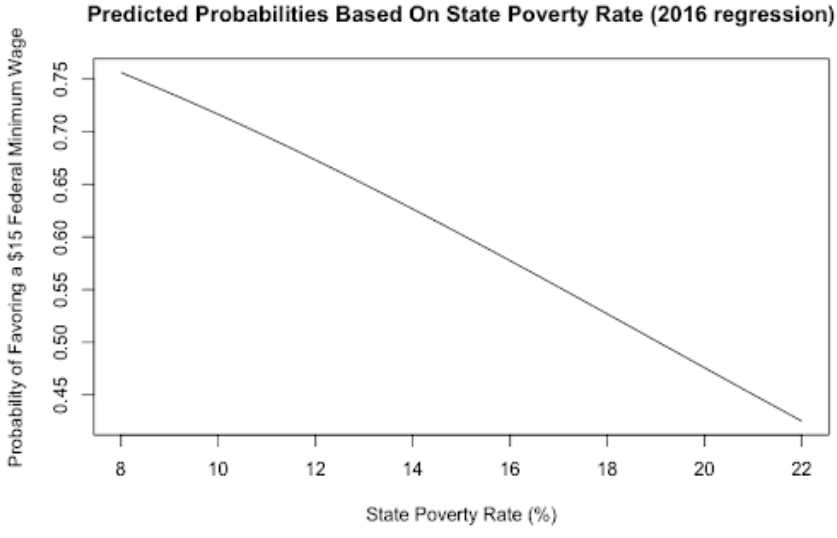
VII.v Figure 1: Predicted Probabilities Based on State Poverty Rate (%) (2013 Regression)



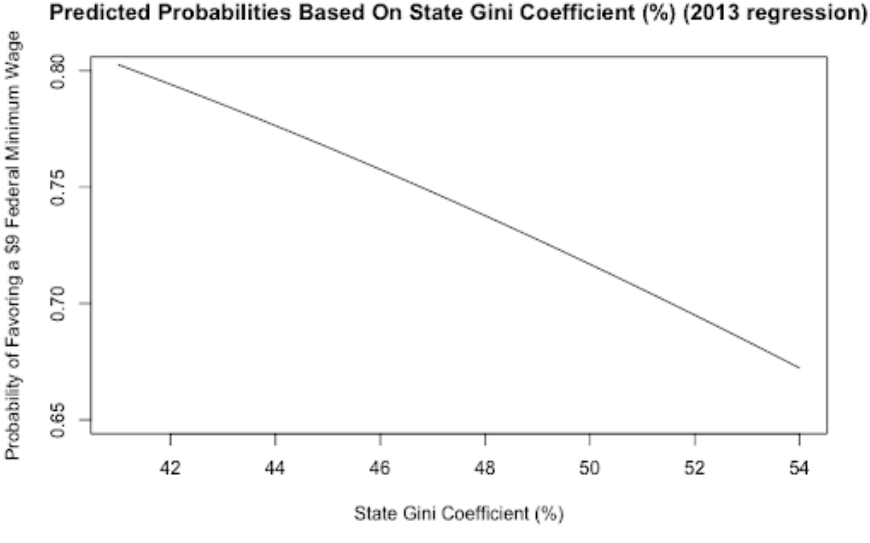
VII.vi Figure 2: Predicted Probabilities Based on State Poverty Rate (%) (2014 Regression)



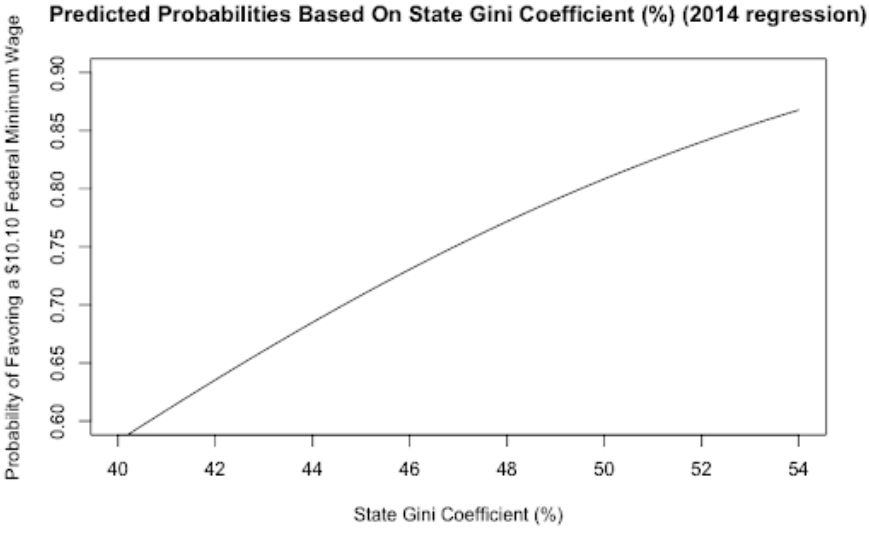
VII.vii Figure 3: Predicted Probabilities Based on State Poverty Rate (2016 Regression)



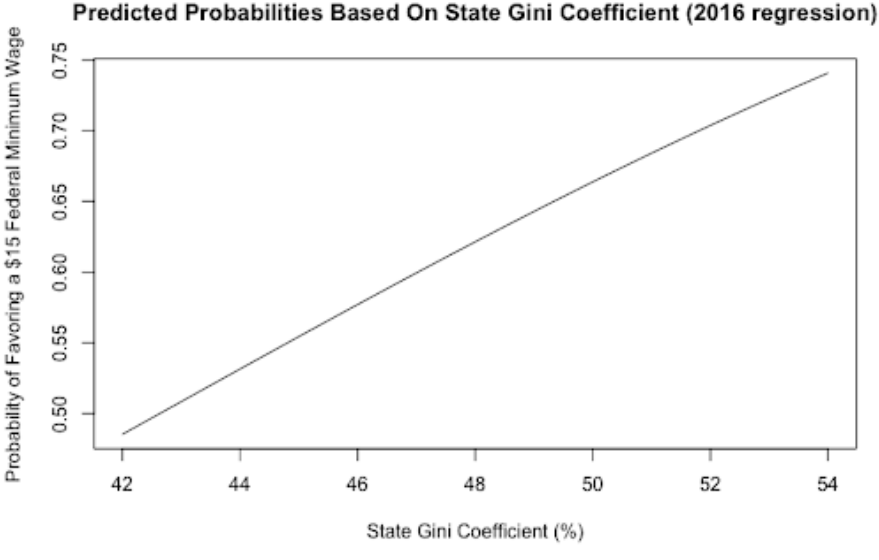
VII.viii Figure 4: Predicted Probabilities Based on State Gini Coefficient (%) (2013 Regression)



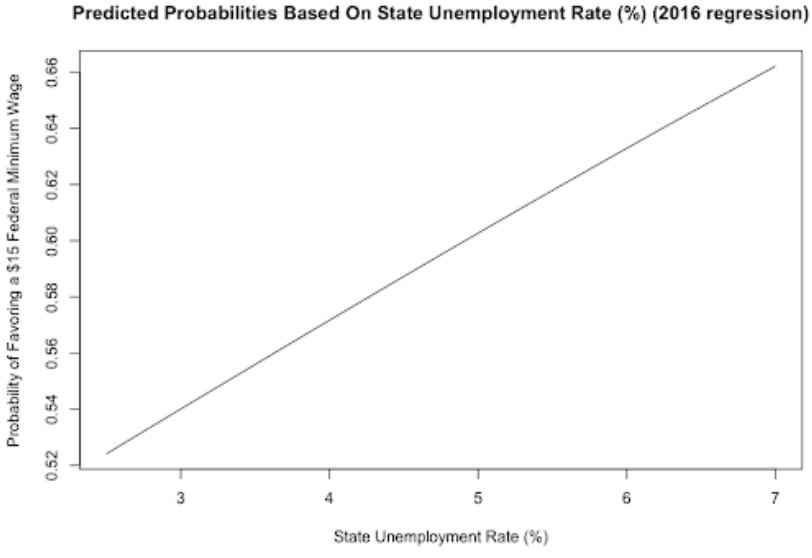
VII.iX Figure 5: Predicted Probabilities Based on State Gini Coefficient (2014 Regression)



VII.X Figure 6: Predicted Probabilities Based on State Gini Coefficient (2016 Regression)



VII.Xi Figure 7: Predicted Probabilities Based on State Unemployment Rate (%) (2016 Regression)



Berkeley
Economic
Review