THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 1 Capitole

Présentée et soutenue par

Kevin REMMY

Le 28 juin 2021

Essays in Empirical Industrial Organization

Ecole doctorale : **TSE - Toulouse Sciences Economiques**

Spécialité : **Sciences Economiques - Toulouse**

Unité de recherche :
UMR 5314 -TSE-R

Thèse dirigée par
Christian BONTEMPS

Jury
M. Steven BERRY, Rapporteur
Mme Rachel GRIFFITH, Rapporteur
Mme Natalia FABRA, Examinatrice
M. Mathias REYNAERT, Examinateur
M. Pierre DUBOIS, Président du jury
M. Christian BONTEMPS, Directeur de thèse

Essays in Empirical Industrial Organization

Ph.D Thesis

Kevin Remmy

Toulouse School of Economics

June 2021

Acknowledgements

The last seven years that I have spent at Toulouse School of Economics have been an incredible journey full of many amazing experiences, tough challenges, and valuable lessons. This thesis is the outcome of this journey. It would not have been possible without the support and guidance of many people to whom I would like to offer my gratitude.

First, I am deeply indebted to my advisor, Christian Bontemps. Christian was there to offer his guidance and help at any moment during the Ph.D. Thanks to him, I could navigate the Ph.D. with the feeling of having someone behind me ready to give support whenever needed. I cannot explain how important this was and how thankful I am for that. It takes more than being a great economist to be the kind of advisor Christian was to me. In my M2 year, he co-offered a master thesis topic for someone to work on. Deciding to pick that topic was probably one of the most important decisions I have made during my doctoral studies. It was undoubtedly the best decision I have made.

I am similarly indebted to my co-advisor, Mathias Reynaert. Mathias has amazed me time and again with his ability to turn my messy ideas into clear and well-defined thoughts. I admire his intellectual rigor and his ability to always find a positive way forward in every problem. His guidance, support, and encouragement were invaluable in getting my job market paper to the point it is. I am also deeply grateful for his vigorous support during the job market. I am incredibly thankful to have had Christian and Mathias as my advisors.

I would like to thank Pierre Dubois for his support over the last year of the Ph.D. and especially during the job market. I am also indebted to Isis Durrmeyer, who has offered guidance and advice on my work since the beginning of the Ph.D. She had to suffer through many poorly written drafts she had to discuss during Ph.D. workshops and always offered invaluable feedback, which helped improve the chapters of this thesis a great deal. Moreover, she and Mathias offered me funding for the last months of the Ph.D., for which I am also grateful. My gratitude also goes to Cristina Gualdani, who has been an amazing co-author and a source of inspiration in many aspects of being a researcher. I have learned a great deal throughout many conversations, not least how to write good slides. I have also benefited from the advice and support of Daniel Ershov, Renato Gomes, Nour Meddahi, Christian Hellwig, Pascal Lavergne, Jacques Crémer, Stefan Ambec, Olivier de Groote, Shruti Sinha, and Nicolas Werquin. The members of the IO and Econometrics Groups at TSE have also provided me with helpful feedback in numerous presentations. I would like to thank Elie Tamer for hosting me at Harvard University and allowing me to spend three months in a highly stimulating research environment. I would also like to thank Rachel Griffith, Steven Berry, and Natalia Fabra for agreeing to be on my committee and providing me with valuable feedback. I am grateful to Skerdilajda Zanaj, who first showed me a career in research was possible and supported me when I decided to take that path.

I am grateful to the members of the TSE support staff, among them Louise Strachan, Laurence Delorme, Ludmila Namolovan, Elvire Jalran, Caroline Tejedor, and Alexandre Noceto, for taking so many things off the Ph.D. students' backs.

Last but certainly not least, I would like to thank my family and friends. Thanks, Philipp, Jeff, Miguel, Elia, Paloma, Rossi, Jasmin, Charles, Mapis, Luise, Romain, Roberto, Anca, Jad, Gosia, Vatsala, Alfonso, and Jannik, for having made Ph.D. life so much more enjoyable. Thanks, Mario, Maxi, Sven, and Steven, for making me feel like I never left whenever I was home. Thanks Mark for the fun times and successful construction projects during the lockdown. Thanks to my "American family" Kirk, Laura, Jack, Audrey, Jemma, and Luke. I am deeply thankful to my siblings Pascal, Michelle, and Nadine, my parents Birgit and Ludwig, and my grandmother Marianne. Their unwavering love and support mean the world to me. Especially, thank you to my fiancee María. Without you, I would never have been able to do this. This thesis is as much your accomplishment as it is mine. I look forward to our road ahead.
I dedicate this thesis to my grandfather.

Abstract - English

This thesis is composed of three chapters. The first chapter studies the economic effects of subsidies when firms can adjust both prices and a product attribute. The second chapter (joint with Christian Bontemps and Cristina Gualdani) builds and estimates a 2-stage model of airline competition. The third chapter (joint with Charles Pébereau) studies adoption of real-time pricing electricity tariffs.

In the first chapter I study the economic effects of electric car subsidies. Electric cars are subsidized around the world because they are seen as a key driver for decarbonizing the transport sector. In response to these subsidies, producers of electric cars can adjust the price of the car, but also the driving range. My analysis finds that subsidy design has an important impact on price and range choices. In the paper, I find that firms react to subsidy directly based on range by selling more expensive electric cars with a higher range. To the contrary, a flat subsidy leads firms to sell cheaper electric cars with a lower range. These findings have important implications for policymakers who have two objectives: Whereas electric car sales are maximized at the flat scheme, minimizing CO2 emissions entails a trade-off between maximizing electric car sales and generating substitution from more polluting cars that is solved at a scheme with a flat and a range-based part. However, whereas a policymaker cannot maximize electric car sales and minimize CO2 emissions with the same scheme, she is able to maximize electric car sales and the consumer surplus of lower-income consumers with the same scheme.

In the second chapter, joint with Christian Bontemps and Cristina Gualdani, we build and estimate a 2-stage model of airline competition. In the model, airlines choose the network of segments to serve in the first stage before competing in prices in the second stage. The two-stage framework allows us to account for selection of airlines into interdependent routes. Moreover, it permits us to make counterfactual exercises that robustly predict changes not only in prices and markups but also in how airlines adjust their route networks. We show that large hub-and-spoke operations lower marginal costs but increase fixed costs. We evaluate a merger between American Airlines and US Airways and compare it to the bankruptcy and disappearance of American Airlines. We also evaluate remedies imposed on the merging parties and find evidence that they limited harm to consumers.

In the third chapter, Charles Pébereau and I study the introduction of real-time electricity pricing in New Zealand and shed light on why adoption was low. Under this tariff, consumers are exposed to half-hourly varying spot prices. We find that prospective and recent adopters

are highly sensitive to contemporaneous spot prices. Adoption rates significantly decrease with contemporaneous spot prices. During a crisis on the electricity spot market, the share of consumers discarding real-time pricing plans decreased with experience. These results suggest that, over time, consumers focus less on immediate outcomes. Our results can inform the debate regarding ways to foster the adoption of real-time pricing, such as opt-in and opt-out policies, and information provision.

Abstract - French

Cette thèse est composée de trois chapitres. Le premier chapitre étudie les effets d'une subvention dans le cas où des entreprises peuvent changer le prix ainsi que les caractéristiques du produit qu'elles proposent. Le deuxième chapitre (joint avec Christian Bontemps et Cristina Gualdani) construit et estime un modèle en deux étapes de la compétition en réseau des compagnies aériennes. Le troisième chapitre (joint avec Charles Pébereau) étudie l'adoption de tarification de l'électricité en temps réel.

Dans le premier chapitre, j'étudie les effets économiques des subventions sur le marché des voitures électriques, qui sont subventionnées partout dans le monde parce qu'on les considère comme un élément clé dans la dé-carbonisation du secteur du transport. En réponse à ces subventions, les producteurs de voitures électriques ont la possibilité d'ajuster le prix des voitures, mais aussi leur autonomie. Mon analyse montre que la façon dont ces subventions sont implémentées a un impact important sur ces choix de prix et d'autonomie. Dans mon étude, je trouve qu'une subvention directement indexée sur l'autonomie conduit les firmes à vendre des voitures électriques qui sont plus chères et qui ont plus d'autonomie. Au contraire, une subvention fixe conduit les firmes à vendre des voitures électriques moins chères avec une autonomie plus faible. Ces effets ont des implications importantes pour les décideurs politiques qui ont deux objectifs: Tandis que la quantité de voitures électriques vendue est maximisée avec la subvention fixe, la minimisation des émissions de CO_2 nécessite un compromis entre maximisation des ventes de voitures électriques et substitution de voitures conventionelles polluant beaucoup. Ce compromis est résolu avec une subvention intermédiaire. Si un décideur politique ne peut pas atteindre les deux objectifs avec la même subvention, il peut tout à fait maximiser les ventes de voitures électriques et le surplus des consommateurs les plus pauvres avec la même subvention.

Dans le deuxième chapitre, écrit en collaboration avec Christian Bontemps et Cristina Gualdani, nous construisons et estimons un modèle en deux étapes du secteur du transport aérien. Dans ce modèle, les compagnies aériennes choisissent d'abord leur réseau de vols directs avant d'entrer en compétition avec ses rivales. Ce jeu en deux étapes nous permet de prendre en compte l'interdépendance des routes choisies par les firmes. En outre, le modèle nous permet de faire des analyses contrefactuelles capables de fournir des prévisions robustes en ce qui concerne le niveau des prix, mais aussi le changement du réseau des firmes. Nous montrons que des réseaux en étoile baissent le coût marginal et augmentent le coût fixe. Nous évaluons une fusion entre American Airlines et US Airways et la comparons au scénario d'une faillite d'American Airlines. Nous évaluons aussi des contre-mesures imposées aux firmes fusionnantes

et trouvons que ces mesures ont réussi à limiter la perte de surplus du consommateur.

Dans le troisième chapitre, Charles Pébereau et moi étudions l'introduction d'une tarification de l'électricité en temps réel en Nouvelle-Zélande et offrons des explications quant à la faible adoption de cette tarication. Cette tarification expose les consommateurs au prix courant de l'électricité changeant chaque demi-heure. Nous trouvons que les consommateurs qui ont adopté cette technologie le plus récemment sont très sensibles aux prix courants. Les taux d'adoption baissent fortement quand les prix courants sont élevés. Durant une crise des prix courants de l'électricité, la part des consommateurs abandonnant la tarification en temps réel baisse avec leur expérience. Ces résultats suggèrent que, au fil du temps, les consommateurs sont moins réactifs aux changement immédiats de prix. Nos résultats sont utiles dans le débat sur la façon d'encourager les consommateurs à adopter une tarification en temps réel, comme des programmes opt-in ou opt-out.

Contents

Chapter 1

Subsidy design when firms can adjust product attributes: The case of electric vehicles

1.1 Introduction

Designing a subsidy that achieves its goal requires knowledge of how firms react to the subsidy. The economics literature has extensively studied how firms with market power react to subsidies when firms can only adjust prices (Bulow and Pfleiderer, 1983; Stern, 1987; Weyl and Fabinger, 2013). In this case, we know the direction of the adjustment: firms respond by lowering prices. The magnitude of the reaction depends on the slopes of the demand and marginal revenue curves. When firms can also adjust product attributes in response to a subsidy, the directions of the price and attribute adjustments are unclear. There exists no evidence in the economics literature on how firms in differentiated product markets adjust prices and product attributes in response to a subsidy.

Knowing what happens when responses by firms can be multi-dimensional is crucial for designing electric vehicle subsidies. Electric vehicles (EVs) play an integral role in reducing carbon dioxide emissions from the transportation sector, and thus countries have committed substantial amounts of funding for EV subsidies. In 2018 alone, world-wide government spending on EV purchases through subsidies totaled \$15 billion.[1] The design of these subsidies differs across countries, with some basing subsidies on product attributes and others granting a subsidy with the same amount to every EV. Policymakers have several objectives in mind when designing these subsidies: minimizing CO2 emissions from new car sales, maximizing diffusion, and attending to distributional concerns. One feature of EVs is that firms can adjust the driving range relatively easily, giving them an additional dimension along which to react to subsidies.[2] The primary determinant of the range is the size of the battery pack and the propulsion mechanism of an EV. Another feature of EVs is that prices of lithium-ion cells, an essential input for battery packs, have dropped substantially over the past decade. This decrease has made the battery pack less expensive and has decreased the cost of providing range.

In this paper, my research question is as follows: How does subsidy design affect market outcomes when firms can adjust prices and product attributes? I use a novel state-level data set of new car purchases in Germany to estimate a structural model of demand and supply. The model allows for flexible substitution patterns across cars of different engine types on the demand side and endogenous range choices by firms on the supply side. I use the estimated model to assess a rich set of counterfactuals. First, I evaluate the impact of lower battery pack prices on the marginal cost of providing range and market outcomes. The supply estimates show that the marginal cost of providing range decreased by 33% from 2012 to 2018 and led firms to sell EVs at higher prices and with a greater range. Firms also collected a higher markup on these EVs. Second, I evaluate a subsidy scheme introduced in Germany in 2018 and find

[1] Source: International Energy Agency.

[2] The driving range, or range henceforth, is the distance that can be driven with a fully charged battery (or, in the case of combustion cars, with a full tank).

that it led to lower prices and a smaller range for electric vehicles on which firms collected a lower markup. Third, I evaluate the effect of different subsidy schemes on market outcomes. Purely range-based subsidies increase price and range, whereas flat subsidies decrease price and range. Policymakers face a trade-off in maximizing diffusion, minimizing CO2 emissions, but can address distributional aspects. Consumers prefer schemes purely based on the driving range of EVs, even though this result hides important distributional effects. Maximizing diffusion and minimizing CO2 emissions from new cars are not equivalent because achieving these two targets requires different substitution patterns. A policymaker can design subsidies that attain one of the objectives or achieve some combination of higher consumer surplus, lower fleet emissions, and greater diffusion.

Several challenges exist in analysing how a multi-dimensional reaction to subsidies and marginal cost changes affects outcomes. First, there exists little guidance in the existing literature on the effect of subsidies in multi-product oligopolies when firms can adjust prices and product attributes. Second, answering this question in the electric car market requires a demand model with rich substitution patterns between electric cars and combustion cars[3], given that the goal of electric vehicle subsidies is to generate more substitution towards electric vehicles. Third, the supply model should allow firms to react to a subsidy by adjusting not only the price but also the range of electric cars. My framework addresses these challenges. I estimate a structural model of demand and supply for new cars. On the demand side, consumers exhibit heterogeneous preferences for cars with different engine types. On the supply side, firms compete in a static oligopoly in which they set the prices of all their products and the range of their electric cars. In general, this model provides a framework for studying the impact of subsidies and marginal cost changes on the price and an adjustable product attribute in multi-product oligopolies. The model builds on Berry, Levinsohn, and Pakes (1995) and the recent literature studying equilibrium outcomes when firms can adjust one or more product attributes (Fan, 2013; Crawford, Shcherbakov, and Shum, 2019). I estimate the model using the generalized method of moments (GMM), using approximations to optimal instruments (Chamberlain, 1987) as proposed by Gandhi and Houde (2019).

Given parameter estimates, I first study the important reduction in prices of lithium-ion cells, a key input for battery packs, which determine the driving range. This input price drop is a defining feature of electric car markets. My framework allows both endogenous provision of range and a multi-dimensional response in terms of price and range to changes in the marginal cost of providing range. I find that the marginal cost of providing range decreased by 33% from 2012 to 2018. Firms pass on this negative shock to the marginal cost of range by selling EVs with a greater range at higher prices. The markup on electric cars increases. These findings are important for subsidy design, as a decrease in the marginal cost of providing

[3]Combustion cars employ a conventional gasoline or diesel engine to propel the car.

range is equivalent to a subsidy purely based on range. Moreover, pass-through occurs through the product attribute channel rather than the price channel. This finding underscores the importance of accounting for a channel through which car manufacturers can adjust range.

In 2016, Germany introduced a subsidy scheme for electric vehicles. The scheme consisted of a flat subsidy, meaning that the amount did not depend on any product attributes. My findings show that the subsidy led to both price and range decreases for electric vehicles, with firms collecting a lower markup. These outcomes are the converse of the adjustment that occurs in response to a lower marginal cost of providing range. In this case, pass-through occurred mainly through the price channel. Prices decreased by more than the amount of the subsidy. Firms used the product attribute channel to reduce range to allow for further price reductions. These two counterfactual exercises bracket the alternative subsidy schemes that I consider in the next step. These alternative schemes consist of a flat part and an incentive-based part that depends on range. The first two counterfactual exercises also make clear how different strategies shape market outcomes: On the one hand, firms can have incentives to increase price and range to target consumers with a high willingness to pay for range and thereby collect a larger markup. On the other hand, firms can have incentives to sell a cheaper product at a lower range, thereby collecting a lower markup but also capturing many consumers with a low willingness to pay for range. Finally, firms can use a subsidy to decrease price and increase range. Which adjustment strategy firms use shapes substitution patterns and, ultimately, market outcomes.

The estimated model allows me to compare a wide range of subsidy schemes and their impact on market outcomes. Across different budgets, I compare schemes that are either flat, are purely dependent on range, or mix a flat part with a range-dependent part. The market outcomes that I focus on are CO2 emissions from new cars and diffusion. In addition, I investigate the effects of subsidy design on consumer surplus and distributional aspects. The ultimate goal of policymakers is to de-carbonize the automobile sector; this makes it natural to look at CO2 emissions from new cars. At the same time, policymakers are interested in increasing diffusion to establish EVs on the market. Dynamic considerations related to learning curve effects also play a role. I find that flat subsidy schemes maximize diffusion. However, maximizing diffusion is not equivalent to minimizing CO2 emissions from new car sales. For the lowest budget considered, CO2 emissions from new cars are lowest at intermediate schemes. Differences in substitution patterns across different subsidy schemes drive this result. Maximizing diffusion warrants a subsidy that maximizes substitution from all cars, whereas minimizing emissions warrants a subsidy that induces more substitution from more-polluting cars. Moreover, I find that the pure range-based scheme maximizes consumer surplus. However, this finding hides substantial heterogeneity: Consumers in lower income deciles prefer purely flat schemes. They do so because the willingness to pay for range decreases with income, meaning consumers at the top of the income distribution have strong preferences for a greater range. In contrast, consumers at the bottom of the income distribution have strong preferences for lower prices. These findings suggest that policymakers can achieve different objectives with different subsidy schemes, with a trade-off in maximizing diffusion, minimizing emissions, and addressing distributional aspects. It is crucial to know how a given scheme affects substitution patterns. Otherwise, the subsidy may have unintended consequences. The results also suggest that a subsidy always decreases emissions, increases diffusion, and raises surplus for every consumer. Accordingly, it is always possible for a policymaker to achieve a combination of lower emissions and higher diffusion while leaving all consumers better off.

This paper contributes to different branches of the literature. The first contribution is to the literature on quality provision (Spence, 1975; Sheshinski, 1976; Mussa and Rosen, 1978; Maskin and Riley, 1984; Crawford et al., 2019; Holland, Mansur, and Yates, 2020) that studies how firms provide a product attribute (quality) in imperfectly competitive markets. The paper also contributes to the pass-through literature (Bulow and Pfleiderer, 1983; Stern, 1987; Kim and Cotterill, 2008; Weyl and Fabinger, 2013) studying how firms adjust prices in response to subsidies, taxes, or marginal cost changes. This paper bridges a gap between these two literature streams by building a framework that allows for a multi-dimensional response in prices and product attributes to subsidies, taxes, and marginal cost changes in imperfectly competitive markets.

This paper also contributes to the literature evaluating environmental policies in car markets. This literature studies how different environmental policies shape market outcomes and compares the effectiveness of different policy tools (Knittel, 2011; Klier and Linn, 2012; Pavan, 2017; Grigolon, Reynaert, and Verboven, 2018; Durrmeyer and Samano, 2018; Reynaert, Forthcoming; Leard, Linn, and Springel, 2019). A sub-branch of this literature is explicitly concerned with the EV market, studying the effect of EV subsidies (Beresteanu and Li, 2011; Xing, Leard, and Li, 2019; Muehlegger and Rapson, 2020) or the impact of charging stations (Li, Tong, Xing, and Zhou, 2017; Li, 2019; Springel, 2020). I contribute to this literature by studying the impact of battery cost changes, one of the defining characteristics of EV markets, on market outcomes. Further, this literature either assumes away supply-side responses of firms or constrains firms to adjust only prices in response to subsidies. This paper contributes to this literature by studying how firms adjust the range in response to subsidies, allowing for an explicit range adjustment channel in the supply model. Finally, I contribute to the literature studying subsidy schemes in EV markets by providing a detailed analysis of how different subsidy schemes affect firm strategies and policy objectives, giving rise to important trade-offs not considered in the literature so far.

1.2 Industry Description and Data

The setting for the empirical analysis is the new car market in Germany. A predominance of combustion engine cars using gasoline or diesel as fuel has characterized this market over the past decades. Simultaneously, sales of electric vehicles increased more than twenty-fold between 2012 and 2018. I estimate both consumer demand for new cars and competition in price and range among firms using a detailed data set of new car transactions.

Industry description

The market for electric vehicles. After having been dormant for more than 100 years, electric vehicle technology came back to prominence in the late 1990s. Both the Honda Insight and the Toyota Prius used a hybrid engine that combined fuel and electric powertrains. However, it was not possible to plug in this electric engine to an external source. Over the past decades, two new technologies have emerged. One is the plug-in hybrid electric vehicle (PHEV), which combines a fuel engine with an electric battery pack that can be plugged into an external power source. The other is a pure battery electric vehicle (BEV), whose powertrain unit consists only of a battery pack (throughout the remainder of the text, "BEV" is used synonymously with "battery electric vehicle", "PHEV" is used synonymously with "plug-in hybrid electric vehicle" and "EV" means both "BEV" and "PHEV"). Electric vehicles have been singled out by policymakers and firms alike as key technologies to de-carbonize the transportation sector in pursuit of the goal to contain the rise of global temperatures to below 2°C. To buttress diffusion, governments around the world have introduced subsidies and tax incentives for electric vehicles. The scope and design of these subsidies vary considerably across and sometimes even within countries. Some countries use flat subsidies, and others make subsidies depend on characteristics such as the driving range or battery size.[4]. Global government spending on EVs increased substantially from \$1 billion in 2012 to \$15 billion in 2018.

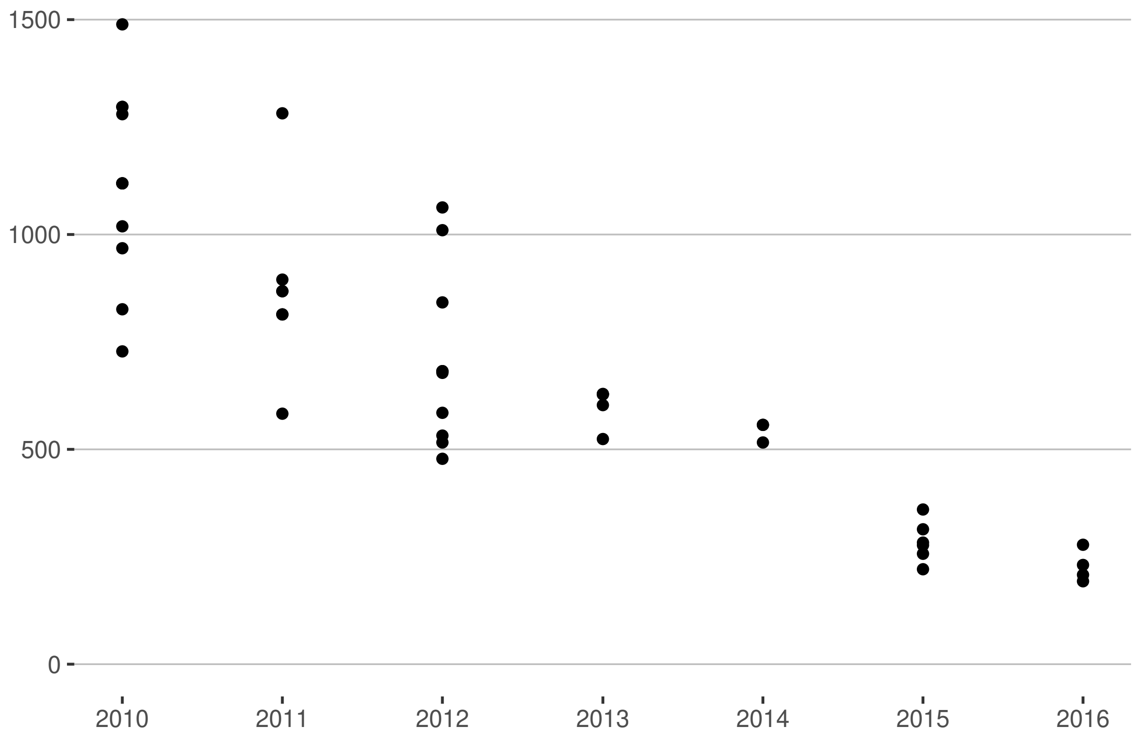
Another feature of the electric vehicle market is the rapid decrease in the cost of lithium-ion cells (LICs). Numerous LICs make up the battery pack of an electric vehicle. This battery pack propels the car, and its size is the most important determinant of the driving range. Figure 1.1 shows different approximations of the evolution of lithium-ion cell prices. Although there is considerable variation in the estimates, there is a clear downward trend. This trend suggests that providing driving range has become considerably cheaper over the past decade.

Significant barriers to the mass adoption of electric vehicles exist: EVs tend to be more expensive and have a shorter driving range than combustion engine cars. In consumer surveys, the high cost and small range of EVs repeatedly show up as the most critical determinants

[4]For detailed overviews, see Yang, Slowik, Lutsey, and Searle (2016) and Rokadiya and Yang (2019).

Figure 1.1: LIC price estimates (USD per kWh)

Source: Hsieh et al. (2019)



of whether to purchase an electric vehicle, together with the charging station network density (see, for instance, Schoettle and Sivak 2018; Carley, Krause, Lane, and Graham 2013; Rezvani, Jansson, and Bodin 2015).

Electric vehicles in Germany. The automobile sector is a key industry in Germany, accounting for 9.8% of gross value added and employing approximately 880,000 people, with another 900,000 jobs heavily depending on the sector, for a combined share of 7.2% of total employment.[5] Germany is home to three of the largest 15 car manufacturers in the world as measured in sales and is ranked fourth in the world in terms of motor vehicle production.

Over the past decade, the German government has implemented measures to boost sales of electric vehicles. One such measure was the Government Program for Electric Mobility of 2016. Part of this program was a support scheme that gave a subsidy of €2,000 for the purchase of battery electric vehicles and a subsidy of €1,500 for the purchase of plug-in hybrid electric vehicles. The car had to have a list price below €60,000 to be eligible for the subsidy. In total, the government provided €600 million in subsidies.[6] The plan reinforced the govern-

[5] https://www.iwkoeln.de/en/studies/iw-reports/beitrag/thomas-puls-manuel-fritsch-the-importance-of-html

[6] Carmanufacturerspledgedtomatchthegovernmentsubsidybygrantingarebateequaltotheamountofthesubsidy. Theprogramalsoprovidedfundingfornewchargingstationsandvarioustaxbenefitsforbuying, using,andchargingelectricvehicles.https://www.bmwi.de/Redaktion/EN/Artikel/Industry/regulatory-environment-and-incentives-for-using-electric-vehicles.html

ment's goal to have 1 million electric cars on the streets by 2020 and 6 million by 2030.[7] The budget was forecast to be sufficient to give subsidies until 2019. However, by June 2017, only approximately 5% of the total budget had been used, and in 2018, the market share of battery electric vehicles was only at 1.2%, with approximately 34,000 annual car sales. These lacklustre sales numbers led the government to increase the subsidy scheme's scope as part of a federal climate protection act in 2019. This act increased the government subsidy for battery electric vehicles to up to €3,000, depending on the list price. The act also increased tax incentives for electric vehicles and introduced a price of €10 per ton on CO_2 from 2021 onward. In total, the government pledged €9 billion for subsidies, tax reductions, and charging infrastructure. Finally, in response to the economic crisis caused by the COVID-19 pandemic, the government doubled the subsidies to €6,000.

Data

I build a comprehensive data set of new car purchases in Germany from 2012 to 2018. I do so by combining several data sources.

Car registrations. I use publicly available data from the German Federal Motor Transport Authority (KBA). This data set contains yearly new registrations at the state level for every car model.[8] A firm-and-trim identifier ("HSN/TSN") defined at a very granular level identifies a model. It differs by car class, body type, engine type, kilowatts, weight, and the number of doors. I follow the previous literature on demand estimation for car markets in treating new registrations as sales.

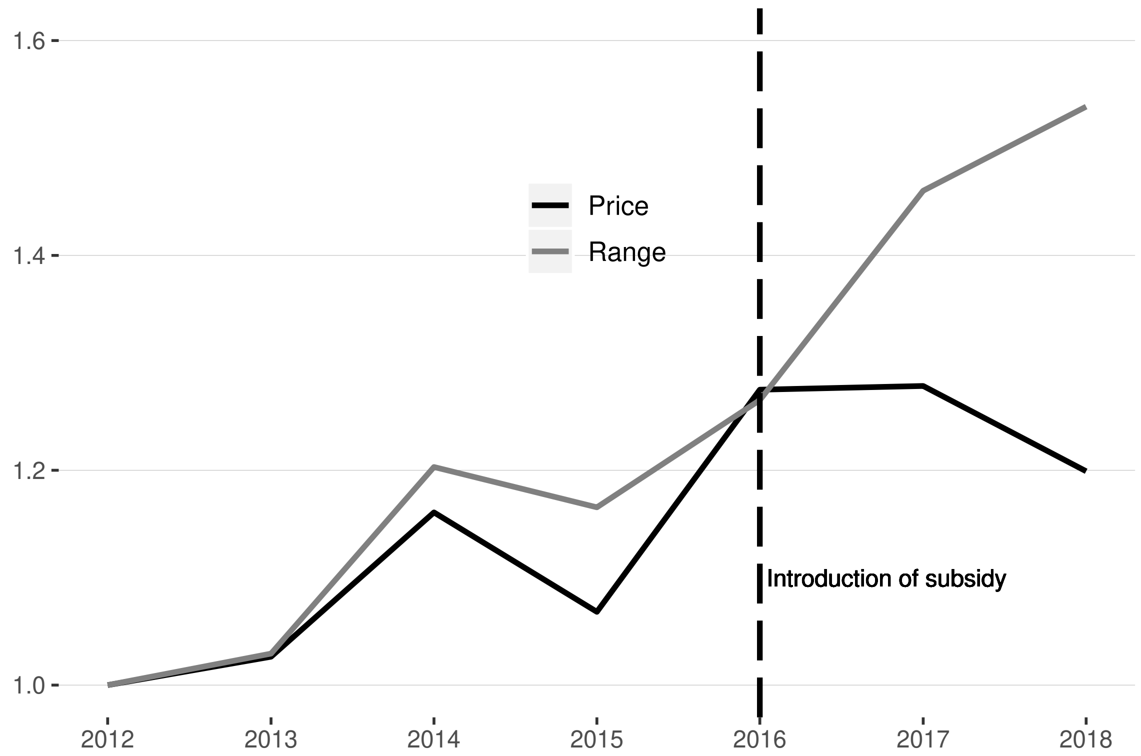
Car prices and characteristics. I scraped data on car prices and characteristics from the website of the General German Automobile Club (ADAC), giving me a comprehensive data set containing a wide range of car characteristics. These characteristics include the driving range of cars. The data also include the list price of cars, which I use in the estimation as the transaction price, again following the literature on demand estimation for car markets. The ADAC data also contain the HSN/TSN identifier, allowing me to match the two data sets relatively easily, except for some observations requiring manual matching.

EV charging stations. I obtain the number of charging stations for electric car batteries from a publicly available data set listing all public charging stations from the Federal Network Agency (BNetzA). The data set contains each station's opening date and its location. This data set allows me to build a variable counting the number of public charging stations in each

[7] https://www.bmwi.de/Redaktion/DE/Downloads/P-R/regierungsprogramm-elektromobilitaet-mai-2011.pdf?__blob=publicationFile&v=6

[8] Germany consists of 16 states ("Bundesländer"). Three of these states (Berlin, Hamburg, and Bremen) are "city-states" whose boundaries coincide with the cities themselves. The other 13 states are larger in area, ranging from approximately the land area of Rhode Island to approximately that of South Carolina. The population of the 16 states ranges from approximately 680,000 (roughly comparable to that of Alaska) to approximately 18 million (roughly comparable to that of New York state).

Figure 1.2: Evolution of price and range of battery electric vehicles (averages, base = 2012)



state each year. I divide the obtained number of public charging stations per state and year by the state's population.

Demographic data. I use data from the German Socio-Economic Panel (SOEP) to build income distributions at the state-year level. To do so, I fit the mean and variance of a log-normal distribution using the observed household income draws of the SOEP. Additional data on population comes from the Federal Statistics Office, and CPI data are from Federal Reserve Economic Data. I build a measure of fuel cost in €/100 km using yearly average gas price data from ADAC and electricity cost data from the German Economics Ministry.

The resulting data set defines a product at a very detailed level. A trade-off exists between having a very granular product definition and a more aggregated definition for tractability. In my final data set, I define a product at the firm/model/engine type level, with the possible engine types being combustion (ICE), plug-in hybrid (PHEV), or battery electric (BEV) engines (e.g., VW Golf ICE vs. Renault Zoe BEV). In aggregating up to this product definition, I use the price and characteristics of the most frequently sold variant at the national level. I reduce the size of the data further by leaving out firms and models with low sales. I set the size of the potential market equal to the number of households in a given state in a given year. In total, the data consist of 28,288 year-state-product observations.

Figure 1.2 shows how the average price and range of battery electric vehicles developed during the sample period. Prices slightly increased, and the range rose by almost 60%. It is unclear from this picture to what extent falling LIC prices and subsidies drove these trends. Detailed summary statistics can be found in Table 2.1 of Appendix A.

1.3 Empirical Model

Set-up

This section introduces a structural model of demand and supply for new cars under endogenous range choices. The model is close to those of Berry et al. (1995), Fan (2013), and Crawford et al. (2019). I need a model that generates realistic substitution patterns between electric cars and combustion cars on the demand side and that explains how firms choose range in a multi-product oligopoly on the supply side. The model also needs to allow me to study the impact of subsidies and marginal cost changes in imperfectly competitive markets when firms choose the price and a product attribute.

Consumers choose the product maximizing their indirect utility and exhibit heterogeneous preferences over prices and product characteristics on the demand side. The supply side allows firms to compete in terms of price and range. I assume that consumers care only about the driving range of battery and plug-in hybrid electric vehicles and not about the driving range of combustion engine cars. These assumptions mirror evidence from consumer surveys on purchase behaviour and consumer preferences regarding battery electric vehicles. Several consumer surveys have found that driving range is the most critical consideration in the purchase of an electric vehicle, next to the price and charging station availability.[9] Additionally, the driving range of combustion engine cars is sufficiently high, and the network of gas stations is sufficiently dense. Hence, this characteristic does not play a role in consumer purchase decisions or firms' profit maximization problems.

I further assume that firms choose prices and range simultaneously at the national level. The rationale behind this assumption is that a firm can alter the driving range even after it has fixed other characteristics, such as the car's size dimensions. A battery pack is made up of many lithium-ion cells, giving firms the flexibility to scale the battery pack's size up or down. Additionally, firms choose price and range at the national level because list prices and characteristics do not vary across states. Finally, I assume that firms only choose their battery electric vehicles' range. This assumption is partly a consequence of the fact that I assume consumers do not have preferences on the range of combustion engine cars. In addition, I

[9] See, for instance, `https://www.compromisorse.com/upload/noticias/002/2794/accentureelectricvehicle.pdf`. Specifically for Germany, see `https://www.aral.de/content/dam/aral/business-sites/de/global/retail/presse/broschueren/aral-studie-trends-beim-autokauf-2019.pdf`. The latter study (in German) also shows that consumers do not take range into account when deciding on the purchase of a combustion engine car.

assume that firms do not choose the range of plug-in hybrid electric vehicles. I do so, first, because the range of PHEVs did not change much over the sample period and, second, because the technology involved is different.[10]

Demand

A state m observed in year t defines a market. There are $\mathcal{M}_{mt}$ consumers in each market mt. Each consumer i chooses one option j, which is either the outside option $j = 0$ or one of the $j = 1, \ldots, J$ differentiated products. Choosing the outside option means that the consumer buys a used car or does not buy a car at all. Choosing one of the "inside" products means buying a new car. The utility that consumer i enjoys from purchasing one of the products $j = 1, \ldots, J$ is

$$u_{ijmt} = r_{jt}\beta^r - \alpha \frac{p_{jt}}{y_{imt}} + x_{jmt}\beta_i^x + \xi_{jmt} + \varepsilon_{ijmt}, \tag{1.1}$$

where r_{jt} is the range of product j, p_{jt} is its price, y_{imt} is the income of consumer i, and x_{jmt} is a vector of observed product characteristics. ξ_{jmt} is an unobserved characteristic of product j in market mt, and ε_{ijmt} is a consumer-specific unobserved taste shock assumed to be an i.i.d. type-I extreme value. The parameter vector β_i^x consists of mean tastes for characteristics and, for some characteristics, random coefficients capturing unobserved heterogeneity in the valuation of product characteristics. For a characteristic k, we have $\beta_i^k = \beta^k + \sigma^k \nu_i^k$ with ν_i^k drawn randomly from a standard normal distribution and σ^k being the standard deviation of preferences. The parameter β^r captures preferences for range, and α captures price sensitivity. Remember that consumers only care about the range of electric vehicles. In the model, this translates into setting $r_{jt} = 0$ for products with a combustion engine. The utility from purchasing the outside option is normalized to $u_{i0mt} = \varepsilon_{i0mt}$.

Consumer i in market mt chooses alternative $j = 0, \ldots J$ that maximizes her utility. Each consumer is characterized by her income y_i and her vector of idiosyncratic preferences ν_i. Income y_i follows a log-normal distribution whose parameters I estimate based on draws from the observed income distribution. Remember that ε_{ijmt} follows a type-I extreme value distribution. This assumption enables me to derive the probability that product j yields the highest utility across all possible alternatives by integrating over the individual-specific valuations for characteristics:

$$s_{jmt}(p, r, x, \xi; \sigma) = \int \frac{\exp(\delta_{jmt} + \mu_{ijmt}(p_{jt}, r_{jt}, x_{jmt}, \xi_{jmt}; \sigma))}{1 + \sum_{k=1}^{J} \exp(\delta_{kmt} + \mu_{ikmt}(p_{kt}, r_{kt}, x_{kmt}, \xi_{kmt}; \sigma))} dF(\nu) dG(y),$$

[10] The battery of a PHEV needs to work in conjunction with a combustion engine. This set-up means that on the one hand, there is less need to increase the range since the combustion engine provides enough range. On the other hand, it is also more difficult to increase the range, given that there are more space constraints.

where $F(\cdot)$ is the joint CDF of the unobserved taste shocks and $G(\cdot)$ is the distribution of income. Further, δ_{jmt} is the mean utility incorporating all terms from (1.1) that do not vary across individuals, and $\mu_{ijmt} = -\alpha \frac{p_{jt}}{y_{imt}} + \sum_k \sigma^k \nu_i^k x_{jmt}^k$ captures individual deviations from the mean utility. Finally, defining the observed market share as $s_{jmt} = \frac{q_{jmt}}{\mathcal{M}_{mt}}$, with q_{jmt} being the observed quantity of product j in market mt, and stacking observed and predicted market shares into a vector, we obtain the system of equations $s_{mt} = s_{mt}(p, r, x, \xi; \sigma)$ for each market mt.

Supply

I model the profit-maximizing price and range decisions of F multi-product firms for each year t. I assume the product portfolio of firms to be given and that firms have already chosen all product characteristics except for the range of BEVs. Firms then maximize profits by setting the price of all products in their portfolio as well as setting the range of their BEVs at the national level.

The profit in year t is then the weighted sum of profits from each state m, and firm f's profit maximization problem can be written as follows:

$$\max_{p,r} \pi_{ft} \equiv \sum_m \phi_{mt} \sum_{j \in \mathcal{J}_{ft}} \Big(p_{jt} - mc_{jt}(r_{jt}, w_{jt}; \theta_s)\Big) s_{jmt}(p, r, x, \xi; \sigma) \mathcal{M}_{mt}, \tag{1.2}$$

where $\phi_{mt} = \frac{\mathcal{M}_{mt}}{\sum_{m'} \mathcal{M}_{m't}}$ is the weight of state m, $\mathcal{J}_{ft}$ is the product portfolio of firm f, $mc(\cdot)$ is the marginal cost of product j, w_j is a vector of observed cost-shifters and θ_s is a vector of parameters entering the marginal cost function. The first-order conditions with respect to price and range are then given by

$$\frac{\partial \pi_{ft}}{\partial p_{jt}} = \sum_m \phi_{mt} \Big\{ s_{jmt} + \sum_{k \in \mathcal{J}_{ft}} \Big(p_{kt} - mc_{kt}\Big) \frac{\partial s_{kmt}}{\partial p_{jt}} \Big\} = 0 \tag{1.3}$$

$$\frac{\partial \pi_{ft}}{\partial r_{jt}} = \sum_m \phi_{mt} \Big\{ - \frac{\partial mc_{jt}}{\partial r_{jt}} s_{jmt} + \sum_{k \in \mathcal{J}_{ft}} \Big(p_{kt} - mc_{kt}\Big) \frac{\partial s_{kmt}}{\partial r_{jt}} \Big\} = 0 \tag{1.4}$$

Equation (1.3) is the usual first-order condition with respect to price, where firm f trades off increasing the margin on product j by increasing the price against losing market share due to this price increase, adjusted by the effect of changing j's price on the demand of other products that firm f offers. We can rewrite (1.4) as

$$\sum_m \phi_{mt} \Big\{ \underbrace{-\frac{\partial mc_{jt}}{\partial r_{jt}} s_{jmt}}_{\substack{\text{Change in markup x} \\ \text{market share}}} + \underbrace{\Big(p_{jt} - mc_{jt}\Big) \frac{\partial s_{jmt}}{\partial r_{jt}}}_{\substack{\text{Markup x change} \\ \text{in market share}}} + \underbrace{\sum_{k \neq j, k \in \mathcal{J}_{ft}} \Big(p_{kt} - mc_{kt}\Big) \frac{\partial s_{kmt}}{\partial r_{jt}}}_{\substack{\text{Cannibalization effect} \\ \text{on other products}}} \Big\} = 0$$

When choosing the range, firm f trades off the decrease in the markup from providing more range (intensive margin) against the higher demand arising from this range increase (extensive/switching margin) as well as the cannibalization effect on the other products in firm f's portfolio. Loosely speaking, equilibrium range decreases with a higher marginal cost of range increases (which squeezes the markup) and increases with larger values of the demand semi-elasticity with respect to range (which increases the extensive margin).

The first-order conditions in (1.3) and (1.4) can be expressed in matrix form. I use the index B for battery electric vehicles and I for other vehicles. I let $\mathcal{J}_B, \mathcal{J}_I$ denote the set of either type of vehicle and J_B, J_I the number of either kind of vehicle on the market. I then define the following matrices:

$$\Delta_p: \ J\text{x}J \text{ matrix with entry } k,l = \begin{cases} \sum_m \phi_{mt} \frac{\partial s_{lmt}}{\partial p_{kt}} \text{ if } k,l \in \mathcal{J}_f \\ 0 \text{ otherwise} \end{cases}$$

$$\Delta_r^B: \ J_B\text{x}J_B \text{ matrix with entry } k,l = \begin{cases} \sum_m \phi_{mt} \frac{\partial s_{lmt}}{\partial r_{kt}} \text{ if } k,l \in \mathcal{J}_f \text{ and } k,l \in \mathcal{J}_B \\ 0 \text{ otherwise} \end{cases}$$

$$\Delta_r^I: \ J_B\text{x}J_I \text{ matrix with entry } k,l = \begin{cases} \sum_m \phi_{mt} \frac{\partial s_{lmt}}{\partial r_{kt}} \text{ if } k,l \in \mathcal{J}_f,\ l \in \mathcal{J}_I \text{ and } k \in \mathcal{J}_B \\ 0 \text{ otherwise} \end{cases}$$

The system of first-order conditions can then be expressed as

$$\begin{cases} \mathbf{s} + (\mathbf{p} - \mathbf{mc})\Delta_p = 0 & (1.5) \\ -\dfrac{\partial \mathbf{mc}^B}{\partial \mathbf{r}^B}\mathbf{s} + \Delta_r^B(\mathbf{p}^B - \mathbf{mc}^B) + \Delta_r^I(\mathbf{p}^I - \mathbf{mc}^I) = 0, & (1.6) \end{cases}$$

where $\mathbf{s}$ is the vector of market shares, $\mathbf{p}$ is the vector of prices, $\mathbf{mc}$ is the vector of marginal costs and $\mathbf{r}$ is the vector of range levels.

Marginal cost specification

I specify a marginal cost function that is log-linear. For product j, it is given by

$$\log(mc_{jt}(q_{jt}, w_{jt}; \theta_s)) = \underbrace{w_{jt}\psi + \omega_{jt}}_{\substack{\text{baseline} \\ \text{marginal cost}}} + \underbrace{(\gamma_0 + \gamma_1 t + \eta_{jt})r_{jt}}_{\substack{\text{marginal cost of} \\ \text{providing range}}}, \tag{1.7}$$

where w_{jt} is a vector of observed cost-shifters, ω_{jt} is a cost shock observed by firms but unobserved by the researcher, t is a linear time trend, η_{jt} is a range-specific marginal cost shock observed by firms but unobserved by the researcher, and $\theta_s \equiv (\psi, \gamma_0, \gamma_1)$ is a vector of parameters to be estimated. Note that the second part of (1.7) is zero for products that are not battery electric vehicles since I do not model their range choices. In the case of BEVs, I assume that the marginal cost of providing range depends on an intercept term, a linear time trend allowing for less costly range provision over time, and an unobserved, product-specific component. The exponential nature of fixed costs is in line with the technology facing firms: Increasing range may be achieved by increasing the size of the battery. A kilometer of range becomes more costly at higher range levels. One reason is that the car's dimensions restrict the size of the battery. Additionally, other ways of increasing range, such as achieving a higher energy density of batteries, may also be constrained by technological factors and make provision of range costlier at higher rangey levels.

Having a functional form for marginal costs allows me to express the equilibrium levels of price and range in matrix form. Let $\mathbf{c}_0 \equiv \mathbf{w}'\boldsymbol{\psi} + \boldsymbol{\omega}$ and $\mathbf{c}_1 \equiv (\gamma_0 + \gamma_1 \mathbf{t} + \boldsymbol{\eta})$. Then, the equilibrium price and range are

$$\begin{cases} \mathbf{p} = \mathbf{mc} + \Delta_p^{-1}\mathbf{s} & (1.8) \\ \mathbf{r} = \dfrac{1}{\mathbf{c_1}} \log\left(\dfrac{\Delta_r^B(\mathbf{p}^B - \mathbf{mc}^B) + \Delta_r^I(\mathbf{p}^I - \mathbf{mc}^I)}{\mathbf{s}^B \mathbf{c}_1}\right) - \dfrac{\mathbf{c}_0}{\mathbf{c}_1} & (1.9) \end{cases}$$

We obtain the usual result of the price being equal to marginal cost plus a markup. The expression for range again makes apparent the trade-off in an increase in market share, cannibalization of other products, and a decrease in the margin or vice versa.

Subsidies in the supply model

The supply model above can accommodate subsidies such as that introduced in Germany in 2016. Let p_{jt} be the price paid by consumers and λ_{jt} the subsidy. Then, the price received by firms is $p_{jt} + \lambda_{jt}$. The profit maximization problem of firm f then becomes

$$\max_{p,r} \pi_{ft} \equiv \sum_m \phi_{mt} \sum_{j \in J_{ft}} \Big(p_{jt} + \lambda_{jt} - mc_{jt}(r_{jt}, w_{jt}; \theta_s)\Big) s_{jmt}(p, r, x, \xi; \sigma) \mathcal{M}_{mt}, \tag{1.10}$$

and the system of first-order conditions is now given by

$$\begin{cases} \mathbf{s} + (\mathbf{p} + \boldsymbol{\lambda} - \mathbf{mc})\Delta_p = 0 & (1.11) \\ -\dfrac{\partial \mathbf{mc}}{\partial \mathbf{r}}\mathbf{s} + \Delta_r^B(\mathbf{p}^B + \boldsymbol{\lambda}^B - \mathbf{mc}^B) + \Delta_r^I(\mathbf{p}^I + \boldsymbol{\lambda}^I - \mathbf{mc}^I) = 0, & (1.12) \end{cases}$$

where $\boldsymbol{\lambda}$ is the vector of subsidies. Expression (1.10) also makes apparent that the introduction of a (flat) subsidy is equivalent to a marginal cost decrease of the firm.

1.4 Estimation

Instrumental variables

Demand side

Estimation of the demand side parameters suffers from the well-known endogeneity issue related to price and here also to range: As the demand- and supply-side shocks realize before the price and range choices, price and range may be correlated with these unobservables. The utility function also includes the number of charging stations available to electric vehicles. The charging station network is itself likely to depend on the electric vehicle base, creating an endogeneity issue (Pavan, 2017; Springel, 2020; Li, 2019). Instruments are needed to account for this endogeneity issue. At the same time, instruments also help identify the random coefficients, thus serving a dual role. Recent literature has pointed out that the classic BLP instruments, consisting of simple sums of product characteristics, tend to perform rather poorly (Reynaert and Verboven, 2014; Gandhi and Houde, 2019). This literature suggests finding approximations for optimal instruments as in Chamberlain (1987). In my estimation, I use differentiation IVs as introduced by Gandhi and Houde (2019). The idea is to describe the relative position of each product in the characteristics space. I build three variants of differentiation IVs: a *local* variant that counts products close in characteristic space, a *quadratic* variant that sums squared differences between product characteristics and a *discrete* variant for discrete variables that counts the number of products with the same value for the characteristic:

$$\begin{aligned}
Z_{jt}^{k,\text{local}} &= \sum_{r \in \mathcal{C} \setminus \{j\}} \mathbf{1}\{|d_{jrt}^{k}| < sd(d^{k})\} \\
Z_{jt}^{k,\text{quadratic}} &= \sum_{r \in \mathcal{C} \setminus \{j\}} d_{jrt}^{k2} \\
Z_{jt}^{k,\text{discrete}} &= \sum_{r \in \mathcal{C} \setminus \{j\}} \mathbf{1}\{|d_{jrt}^{k}| = 0\}
\end{aligned}$$

where $|d_{jrt}^{k}|$ is the absolute value of the difference between products j and r in characteristic k, $sd(d^{k})$ is the standard deviation of characteristic k across markets, and $\mathcal{C}$ is the set of products considered. I build four kinds of instruments of each variant: one considering own-firm products, one considering rival-firm products, one considering own-firm products of the same engine type (BEV, PHEV, or ICE) and one considering rival-firm products of the same engine type.

I build the local and quadratic variants for all continuous characteristics and the discrete variant for all discrete characteristics. I also create local and quadratic variants for a price index, obtained from regressing the observed price on demand- and cost-shifters. The range of BEVs is endogenous, but I assume that the range of PHEVs is not. This is why I build the local and quadratic variants for the range of plug-in hybrid vehicles. I also build the local and quadratic variants for battery efficiency (measured in kWh/100 km), which I assume to be exogenous. I use a subset of all the instruments that I create. I account for the endogeneity of the charging station network by including subsidies as instruments. These subsidies vary across years as well as across states.

Supply side

On the supply side, firms choose range after they have fixed all other product attributes. Range choices can thus be correlated with unobserved marginal cost shocks. I account for this endogeneity issue by constructing differentiation IVs built on the exogenous characteristics entering the marginal cost function. I also include the observed exogenous characteristics entering the baseline marginal cost, as these characteristics were chosen before range, guaranteeing their exogeneity with respect to the unobserved range-specific cost shock. As on the demand side, I use a subset of the instruments that I create.

Identification

Using the set of instruments described above allows me to pin down the estimated parameters. I recover the mean utility parameters β and the cost parameters ϕ through a linear projection. Variation in market shares and observed characteristics then identify β. Market share variation exists across states (the m part of the market index) and time (the t part of the market index). In contrast, product characteristics mainly vary across time (except for the endogenous charging station variable). The demand-side parameters, coupled with an assumption on firm behaviour, allow me to back out implied marginal costs. Changes in the implied marginal cost and observed cost-shifters then identify the vector of marginal cost parameters ϕ. In addition to using the instruments described above, variation in the observed characteristics helps identify σ. Similarly, variation in market shares, prices, and consumer income identify the price coefficient α. Prices vary across time, whereas consumer income varies both across time and across states. The parameters (γ_0, γ_1) governing the marginal cost of additional range are identified from variation in observed range levels and the implied marginal cost of providing it, which, in turn, depends on variation in prices and market shares. For a more elaborate discussion on the identification of demand and supply models with differentiated products, refer to Berry and Haile (2014).

Zero market shares

Approximately 4% of my observations are products with strictly positive national-level sales but zero state-level sales. Zero sales pose a problem in random coefficient demand models, as the estimation procedure is not well defined when zero sales are present. Deleting observations with zero sales from the sample is problematic because it alters the market structure and makes these products unavailable in counterfactual analyses. There exist approaches in the literature to accommodating zero sales in random coefficient demand models.[11] I follow D'Haultfœuille, Durrmeyer, and Février (2019) and use a simple correction of market shares:

$$s_j^c = \frac{q_j^{obs} + 0.5}{\mathcal{M}},$$

where q_j^{obs} is the observed quantity sold of product j in a given market and $\mathcal{M}$ is the market size in that market. This correction aims to minimize the bias of $\log(s_j)$ such that demand parameters can be consistently estimated. D'Haultfœuille et al. (2019) provide an interesting and detailed discussion on this.[12]

Estimation of the demand side

On the demand side, the vector of parameters to be estimated is given by $\theta_d \equiv (\beta_i^x, \beta^r, \alpha)$. I allow random coefficients on characteristics for which I believe consumer heterogeneity matters: an *EV* dummy for battery- and plug-in hybrid vehicles and *Fuel Cost*, measured in €/100 km. The random coefficient on the *EV* dummy allows flexible substitution between electric cars and combustion engine cars. Obtaining such flexible substitution patterns is crucial for studying the market outcomes of subsidy schemes, as substitution across engine types drives these outcomes. The random coefficient on *Fuel Cost* allows consumers to have idiosyncratic preferences for a characteristic that proxies the usage cost of cars. Additionally, substantial differences across engine types exist in the fuel cost per 100 km, which renders the substitution patterns between cars of different engine types more flexible. I allow a trend in the mean taste for range, possibly capturing taste changes for range over time. In addition, I add several characteristics for which I only estimate the mean taste, including the number of public charging stations per 10,000 inhabitants, fuel cost, footprint, doors, dummies for electric vehicles, a dummy if the firm has its headquarters in the state considered, and a linear time trend.[13] I also add brand, class, body and state fixed effects. All remaining unexplained variation is then collected in ξ_{jmt}, which is interacted with the instruments described in the previous section to build moment conditions of the form $E[z_{jmt}^d \xi_{jmt}] = 0$, with z_{jmt}^d as an instrument. Stacking ξ_{jmt} across products and

[11] Li (2019) uses a Bayesian shrinkage estimator to move market shares away from zero. Lu, Shi, and Gandhi (2019) construct bounds for the conditional expectation of inverse demand and show that their approach works well even when the fraction of zero sales is 95%. Dubé, Hortaçsu, and Joo (2020) use a pairwise-differencing approach to estimate demand parameters.

[12] The zero sales problem is rather small in my sample, given that it only affects approximately 4% of my observations. My results are robust to the use of different corrections (such as replacing $q_j = 0$ with $q_j = 1$), which I see as evidence that my demand parameters are consistently estimated and lead me to believe that the correction I use is sufficient.

[13] I introduce the last variable to account for the fact that car companies often register a large number of cars in their home state. Firms do so to comply with emissions regulations or to sell these cars at a discount later. Not accounting for this may introduce a bias, especially for products with small market shares.

markets into a column vector ξ, I obtain the GMM objective function to be minimized:

$$\min_{\theta_d} \xi(\theta_d)' Z^d W^d Z^{d'} \xi(\theta_d),$$

where Z^d contains the instruments and W^d is a positive definite weighting matrix. I use the two-step efficient GMM estimator, where I use an approximation of the optimal weighting matrix based on an initial set of estimates to recover the final estimated vector of parameters. The estimation algorithm that I use is described in detail in Berry et al. (1995) and Nevo (2001). In the estimation, I account for various numerical issues that recent literature has drawn attention to (Dubé, Fox, and Su (2012), Knittel and Metaxoglou (2014), Brunner, Heiss, Romahn, and Weiser (2017), Conlon and Gortmaker (Forthcoming)). First, I approximate the market share integral with 1,000 draws using modified Latin hypercube sampling. Hess, Train, and Polak (2006) and Brunner et al. (2017) show that this method performs very well in random coefficient logit models and provides better coverage than the more frequently used Halton sequences. Second, I set the tolerance level in the contraction mapping of the inner loop to 1e-14 to solve for the demand-side unobservables. A tight tolerance prevents numerical errors from the inner loop from propagating to the outer loop. Third, I use the low-storage BFGS algorithm of NLOPT. Fourth, I initialize the optimization routine from many different starting values to search for a global minimum. Finally, I check first- and second-order conditions at the obtained minimum to ensure the optimizer did not get stuck at a saddle point.

Estimation of the supply side

With demand estimates in hand, I can derive implied markups and marginal costs. The vector of parameters to be estimated is $\theta_s = (\psi, \gamma_0, \gamma_1)$. I let the baseline marginal cost depend on several observed characteristics, such as the product's weight, footprint, fuel efficiency, and engine power measured in kilowatts. I also include year, firm, class and body fixed effects. All remaining unobserved marginal cost-shifters are then collected in ω_{jt}.

Remember that the marginal cost of additional range consists of an intercept and a linear time trend to capture the decreasing cost of the lithium-ion cells that are a crucial input for the battery pack, the size of which, in turn, is a main determinant of range. Any unobserved, product-specific cost of additional range is then captured by η_{jt}.

The first-order conditions in (1.5) and (1.6) can be solved for the pair of supply-side unobservable vectors $\boldsymbol{\omega}$ and $\boldsymbol{\eta}$. I then interact them with the instruments described in the previous section to build moment conditions of the form $E[z^s_{jt}\omega_{jt}] = 0$ and $E[z^s_{jt}\eta_{jt}] = 0$. Letting $\rho_{jt} = (\omega_{jt}, \eta_{jt})$ and stacking across products and markets, I then obtain the GMM objective function to be minimized:

$$\min_{\gamma_0,\gamma_1} \rho(\gamma_0, \gamma_1)' Z^s W^s Z^{s'} \rho(\gamma_0, \gamma_1),$$

where Z^s contains the instruments and W^s is a positive definite GMM weighting matrix. The baseline marginal cost parameters ψ can be concentrated out of the minimization routine, much like the linear mean tastes in the utility function. Note that the number of observations differs on the demand and supply sides. As firms choose price and range at the national level, I have one national market per year t and not m state-level markets per year t on the supply side.

I take into account subsidies as outlined in (1.11)-(1.12). I do not consider rebates granted by firms for two reasons: The first is that some firms granted larger rebates than they had pledged. I do not observe these rebates. The second reason is that during the sample period, firms also granted substantial rebates on gasoline and especially diesel cars, to a large extent in response to the Volkswagen emissions scandal.[14] The list prices net of government subsidies can be seen as the maximum transaction price, as is the case in most of the literature estimating demand and supply in new car markets.

1.5 Results

The estimated coefficients of key parameters are in Table 1.1. The first three columns show demand estimates, and the last three columns show marginal cost estimates along with standard errors in parentheses. The full estimation results including fixed effects are in Appendix B. Table A.2 in Appendix A reports the first stage. Overall, the signs and magnitudes of the estimated coefficients are in line with standard economic intuition.

Consumers like greater range, all else equal. The range-specific trend is negative, meaning that consumer preferences for range become less intense throughout the sample period. One explanation for this could be that range anxiety has decreased over time due to consumers learning more about electric vehicles. This learning may come from their own experience, that of peers, or simply a greater availability of information on electric cars. Research and consumer surveys suggest that the driving range of current battery electric cars is sufficient for most trips. Li, Linn, and Muehlegger (2014), for instance, report that households drive approximately 50 miles per day on average. Another explanation may be that faster battery charging has made consumers less worried about range. A further explanation for the negative trend is that it captures decreasing marginal utility of range as the range increases. Such an increase in the range of electric vehicles has indeed occurred, as evidenced in Table 1.2. The positive and

[14] https://www.handelsblatt.com/unternehmen/industrie/studie-zum-automarkt-wo-es-die-groessten-diesel-rabatte-gibt/22682110.html?protected=true

Table 1.1: Key demand and marginal cost estimates

	Utility		Marginal Cost		
	Coefficient	SE		Coefficient	SE
Mean Utility			**Range Provision**		
Range	1.772	(0.223)	Intercept	0.813	(0.026)
Range × Trend	-0.118	(0.024)	Trend	-0.070	(0.006)
Charging Stations	0.610	(0.156)			
Fuel Cost	-0.281	(0.027)			
BEV	-8.285	(1.539)			
PHEV	-5.901	(1.482)			
Interactions					
Price/Income	-5.713	(0.691)			
Standard Dev.					
EV	2.563	(0.685)			
Fuel Cost	0.134	(0.017)			
Statistics					
Mean own-price elasticity	-3.267				
Mean own-range elasticity (BEVs)	2.854				
Mean markup (€1,000)	10.556				

Note: Prices are deflated and in €1,000. Vehicle class, body, firm and state fixed effects included. See Appendix B for the full estimates.

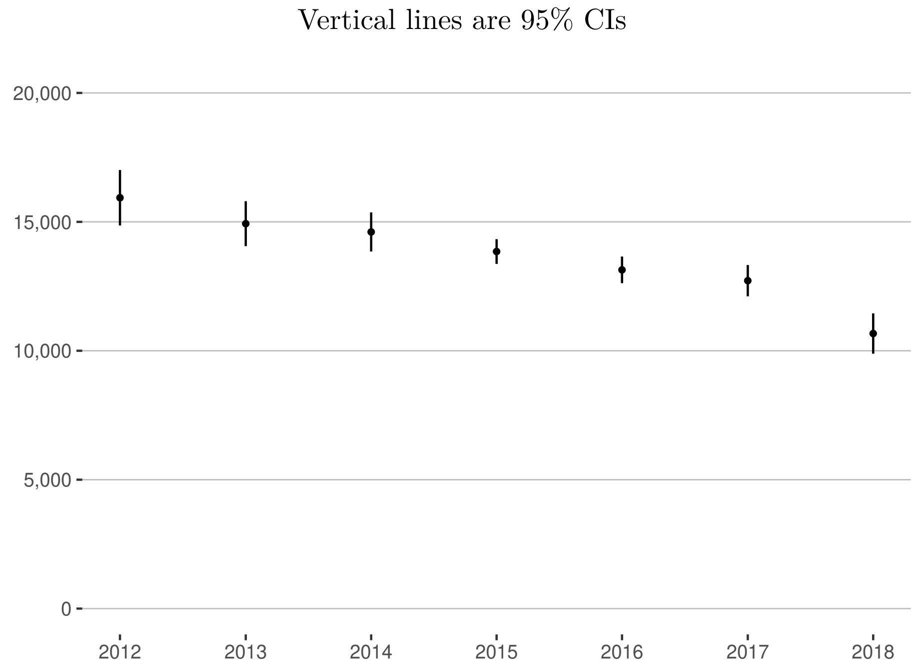
statistically significant sign on the *Charging Station* variable implies that consumers prefer more charging stations, in line with previous studies on demand for electric vehicles (Li, 2019; Springel, 2020). The mean range elasticity is equal to 2.854.

The negative and significant coefficient on price over income translates into a mean price elasticity of -3.267, which falls within the range of figures found in the long literature on demand estimation for new car markets. Table D.1 in Appendix D shows how my estimated price elasticity compares to those found in other papers. Unlike the sensitivity of range, price sensitivity barely changes over the sample period. Due to slightly larger and slightly more dispersed household income, mean price sensitivity dropped slightly from 2012 to 2018, with the variance increasing slightly. Graphical evidence of the findings is provided in Figure A.3 in Appendix A. The relative stability of price sensitivity, together with the finding of a lower valuation of range over time, suggests that towards the end of the sample period, consumers valued (a lower) price more relative to range than at the beginning.

All else equal, consumers strongly dislike both battery and plug-in hybrid electric vehicles, even though there is considerable heterogeneity in the population. A small share of consumers prefer electric cars over those with a combustion engine. The results suggest that the disutility from purchasing EVs decreased over the sample period since the driving range and the number of charging stations increased. This finding also underscores the importance of range and charging stations for the mass adoption of EVs.

Consumers dislike higher fuel costs, as evidenced by the negative parameter in the mean utility. A dis-utility from higher driving costs makes sense, as these increase the overall cost of using a car. However, consumers exhibit considerable heterogeneity in their valuation of fuel costs. Heterogeneity in the valuation of fuel costs is also unsurprising, as factors such as income, driving behaviour, and preferences for less fuel-efficient cars play a role in shaping an individual's fuel cost valuation.

Figure 1.3: Estimated yearly mean marginal cost of providing range



On the marginal cost side, I find that range is costly to provide. Range provision became cheaper over the sample period, evidenced by the trend's negative and statistically significant coefficient. This trend translates into a mean decrease in the marginal cost of providing range of approximately 33% from 2012 to 2018 (see Figure 1.3). This number is somewhat lower than the estimates of lithium-ion cell price decreases in Hsieh et al. (2019), for instance. Given that car manufacturers import most lithium-ion cells from overseas and may not directly benefit from price drops due to long-term contracts, it seems plausible that the fall in the marginal cost of providing range follows the lithium-ion cell price decrease less than one to one. Another explanation is that firms need to convert the lithium-ion cells into a battery pack that is an important – but not the exclusive – determinant of driving range.

Figure 1.4: Estimated marginal cost functions for 2012 and 2018

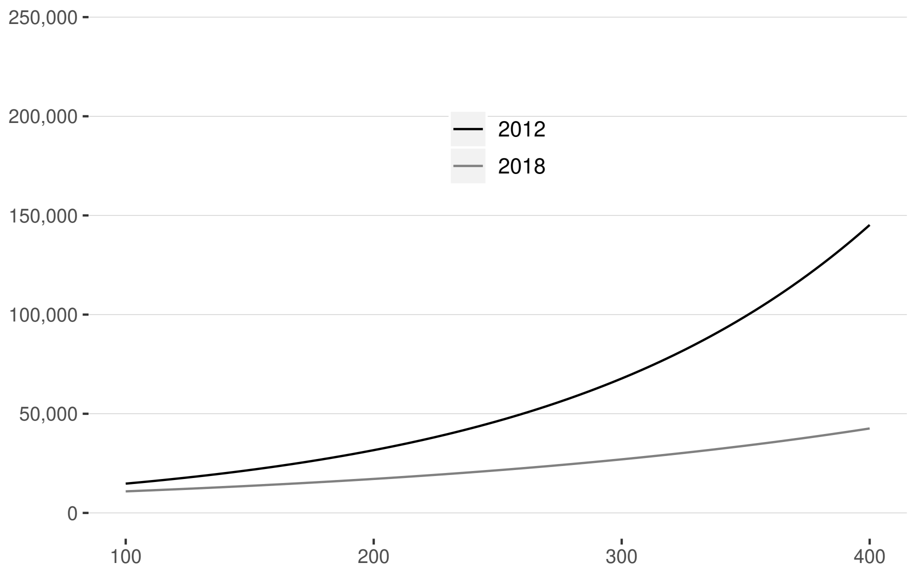
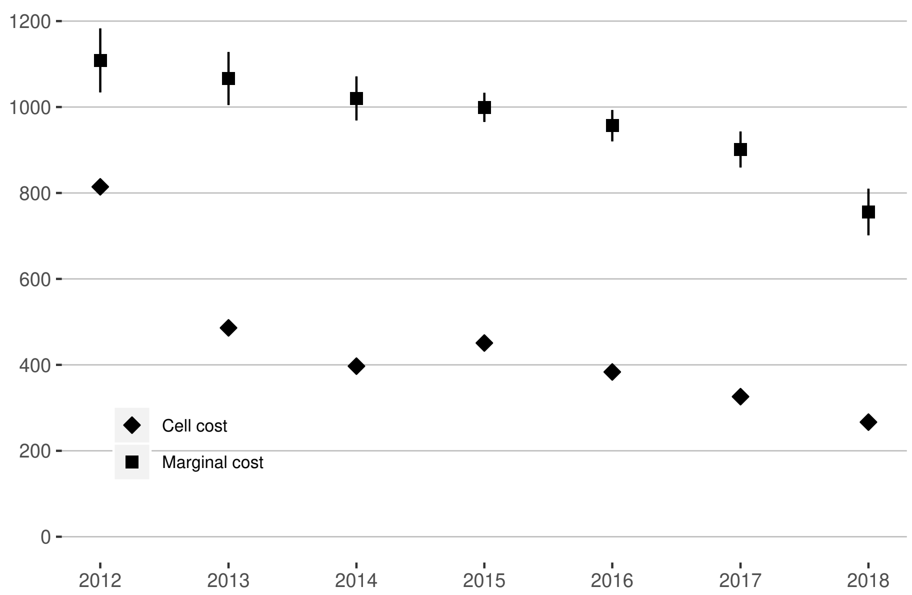


Figure 1.4 plots marginal cost curves at different range levels for 2012 and 2018. The lines are computed using the mean estimated baseline marginal cost across BEVs and the mean estimated marginal cost of providing range for 2012 and 2018, respectively. The curve is much "flatter" in 2018 than in 2012, when range levels higher than 200 km resulted in a marginal cost above €50,000. The figure suggests that it was not feasible to provide many of the range levels observed in 2018 at a competitive price.

Figure 1.5: Per-kWh cost at observed range levels against battery pack cost



To dig deeper into the validity of the marginal cost estimates, I translate the marginal cost of providing range into a battery cost per kWh. Dividing the estimated mean marginal cost of providing range by the battery efficiency, I obtain a cost per kWh. I then compare this per-kWh translation of the marginal cost of providing range to estimated costs of a battery pack, taken from an engineering report (Steen, Lebedeva, Di Persio, and Boon-Brett, 2017). This report provides an estimate for the battery pack cost in $ per kWh for the sample period considered, which I convert into euros and deflate. The results are shown in Figure 1.5. We can see that the estimated per-kWh cost, evaluated at observed range levels, is above the battery pack cost coming from engineering estimates. This finding makes sense, given that the battery pack's size is the main but not the only determinant of providing range. Additionally, the graph shows the per-kWh cost evaluated at observed range levels and imputed marginal cost levels. Given the log-linear marginal cost specification, this per-kWh cost would be different at different marginal cost and range levels. However, apart from 2012, the per-kWh cost backed out of the model follows a similar trend to the battery pack estimate, providing evidence that my marginal cost estimates are reasonable.

The baseline marginal cost estimates have the expected signs and magnitudes. Larger, heavier, more powerful, and more fuel-efficient cars are more costly to produce. Battery electric vehicles are cheaper to produce, all else equal, which is reasonable given that apart from the costly range provision, there are many parts (gearbox, exhaust pipe, starter, injection system, etc.) that are not necessary in the production of a BEV. The supply-side results suggest that range provision accounts for approximately 62% of the marginal cost of producing a BEV, on average. This finding is in line with recent engineering cost estimates (Lutsey and Nicholas, 2019), further suggesting that my marginal cost estimates are reasonable in magnitude.

1.6 Counterfactuals

In this section, I use the estimated model to quantify the effect of marginal cost changes and subsidies on battery electric vehicles by performing several counterfactual exercises. In a first step, I evaluate how firms use a lower marginal cost of providing range to adjust the price and range of their BEVs. In a second step, I assess the subsidy scheme imposed in Germany to see how firms adjusted price and range in response to the subsidy. Finally, I evaluate different subsidy schemes and compare them in terms of market outcomes. This step allows me to describe how subsidy design affects policy objectives and the underlying substitution patterns. It also allows a discussion on the compatibility of different policy objectives.

Procedure

Having estimates of price and range semi-elasticities, a system of first-order conditions (FOCs) for prices and range levels, and an estimate of the marginal cost of providing range, I can compute the new equilibrium vectors of price and range. I employ an iterative algorithm to find the new equilibrium vector of prices and range levels $(\mathbf{p}, \mathbf{r})$. I proceed as follows: At iteration h,

1. Use the price FOCs to compute $\mathbf{p}^{h+1} = \mathbf{mc}(\mathbf{r}^h) + \Delta_p^{-1}\mathbf{s}(\mathbf{p}^h, \mathbf{r}^h) - \boldsymbol{\lambda}$

2. Update market shares and elasticities using $\mathbf{p}^{h+1}, \mathbf{r}^h$

3. Use the range FOCs to compute $\mathbf{r}^{h+1} = f(\mathbf{r}^h, \mathbf{p}^{h+1})$, where $f(\cdot)$ is the expression of range from (1.9).

4. Update market shares and elasticities using $\mathbf{p}^{h+1}, \mathbf{r}^{h+1}$

5. Let $d_{max} = \max(d_p^h, d_r^h)$, where $d_p^h = \max|\mathbf{p}^{h+1} - \mathbf{p}^h|$ and $d_r^h = \max|\mathbf{r}^{h+1} - \mathbf{r}^h|$

6. If $d_{max} \geq \epsilon^c$ with ϵ^c being some convergence criterion, go back to step 1. If $d_{max} < \epsilon^c$, extract $(\mathbf{p}^{h+1}, \mathbf{r}^{h+1})$ to be the new equilibrium vector of prices and range levels.

I adapt the algorithm for counterfactuals in which only price or only range is allowed to be re-adjusted simply by using the respective FOCs only. I find that this procedure converges to the same vector of prices and range levels even when I start from different starting values in different counterfactual settings, which I take as a sign that there exists a unique counterfactual equilibrium. Altering the ordering of the price and range updating does not change the results. The same holds for an alternative procedure, where I iterate until convergence on, say, price in an "inner loop" before iterating until convergence on the range and repeat both iterations until the "outer loop" converges. These alternative procedures give me confidence that the counterfactual results that I find are robust to the specific algorithm and different starting values. The fact that firms choose only the range of BEVs means that the number of additional FOCs to iterate in addition to the price FOCs is small. Additionally, the first-order price changes are confined to BEVs. These factors contribute to the good convergence properties of the algorithms. I perform all counterfactuals for 2018.

How does a lower marginal cost of range provision affect price and range?

On the supply side of my model, I find that the marginal cost of range provision decreased by approximately 33% between 2012 and 2018. Evidence from the engineering literature and from policy reports suggests that a primary driver of this marginal cost drop has been falling lithium-ion cell prices. While there is uncertainty on the future path of these prices, there is a general agreement that they will continue to fall over the next 5-10 years (Hsieh et al., 2019; Nykvist and Nilsson, 2015; Green, Armstrong, Ben-Akiva, Heywood, Knittel, Paltsev, Reimer, Vaishnav, Zhao, Gross, et al., 2019).

Table 1.2: Market outcomes with lower marginal cost of range

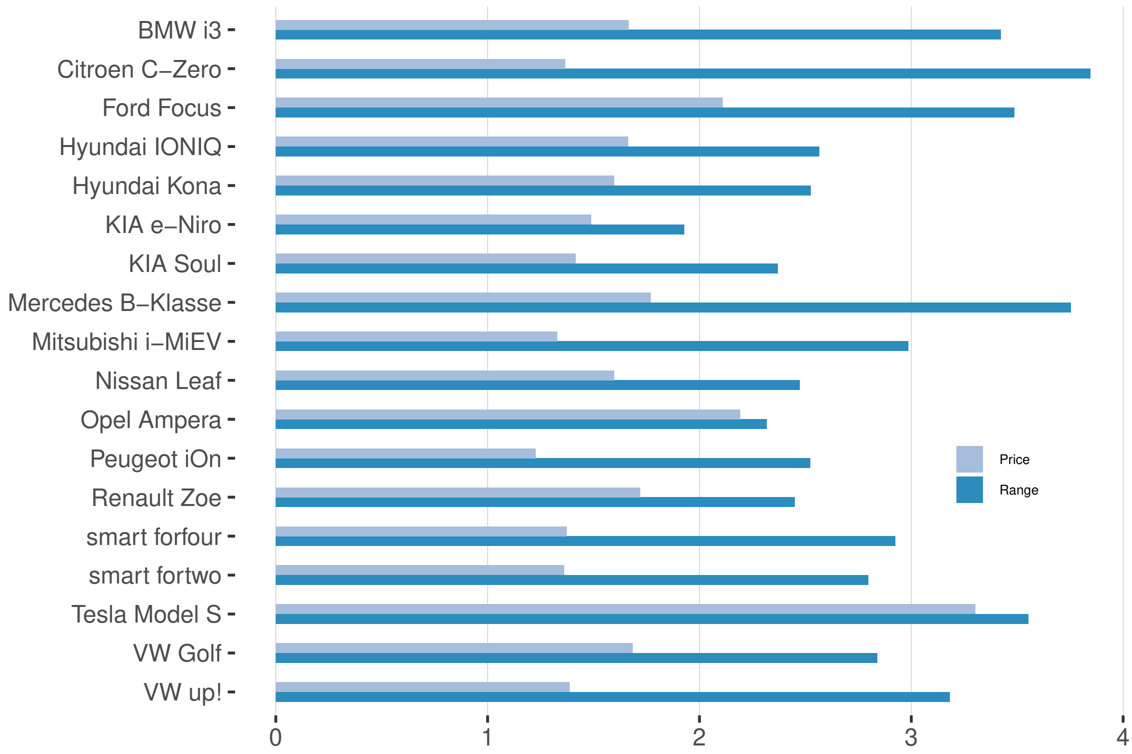
	Observed	**With lower marginal cost**		
	Base	Price, range adjust	Only price adjusts	Only range adjusts
Price	34,671	+659	-357	0
		(+584, +808)	(-468, -226)	
Range	259	+7	0	+3
		(+6, +15)		(+3, +4)
MC	26,023	+554	-301	+59
		(+482, +690)	(-393, -192)	(+54, +65)
Markup	10,536	+104	-56	-59
		(+102, +123)	(-81, -35)	(-65, -54)
Sales	34,761	+594	+1,044	+1,057
		(+290, +811)	(+700, +1313)	(+790, +1287)

Notes: The table gives mean differences from observed outcomes with 95% CIs in parentheses.

In principle, firms can pass through lower marginal costs to price or range. The pass-through rates are given by $\frac{dp}{dc_1}, \frac{dr}{dc_1}$, where c_1 denotes the marginal cost of providing range. The signs and magnitudes of the rates are determined by the relative price and range semi-elasticities, the marginal cost of providing range, cannibalization effects, and the strategic effect of firms' own actions on rival firms. To find the direction and magnitudes of the effects, I re-compute the market equilibrium when the marginal cost of providing range drops by 1%.
I perform this counterfactual under three scenarios. In scenario 1, I allow firms to adjust prices and range. In scenarios 2 and 3, I restrict firms to adjusting price only and adjusting range only, respectively. The results are shown in Table 1.2.

The observed outcomes are in the second column of the table. The third column shows the results when the marginal cost of providing range decreases by 1% and both price and range can be adjusted. We can see that instead of passing through the cost decrease to lower prices, prices increase. The reason for this price increase is that firms improve range. Firms now sell a more expensive product with a higher range. The markup increases as well, suggesting that it is a profit-maximizing strategy to emphasize the intensive margin (charging a higher markup

Figure 1.6: Percentage changes of price and range due to lower marginal cost of range



on existing consumers) over the extensive margin (attracting additional consumers). In other words, the firm finds it more profitable to attract consumers with a high willingness to pay for range as opposed to consumers who care relatively more about price than about range. In the last line of Table 1.2, we see that sales also increase by approximately 1.5%. Figure 1.6 shows that the direction of the effects is uniform across battery electric vehicles sold in 2018.

The third and fourth columns of Table 1.2 present the outcomes when only price and only range can adjust, respectively. In the case of pure price adjustment, the average price decreases by approximately €357 or 1.03%, meaning that there is a slight over-shifting of the marginal cost drop. In the case of pure range adjustment, the average driving range increases by 3 km or 1.15%, also suggesting over-shifting. We can also observe that when firms can only adjust a single variable, the markup decreases, as opposed to the case where both range and price are free to change, suggesting that competitive effects are important. The ability to adjust both price and range allows firms to increase their markup, something that they do not find it profitable to do when they can only adjust price.

This section provides an answer to the question of how a marginal cost shock affects the price and range of battery electric vehicles: A negative marginal cost shock increases both prices and range levels, leading to more expensive products with higher range on which firms collect a higher markup. This result suggests that range can be expected to increase. However, this result depends on the current levels of price, range, and marginal cost of providing that range. When price, range, and the marginal cost of providing range are different, the directions and magnitudes of the effects of a lower marginal cost of providing range on price and range may differ.

How did the German subsidy scheme affect price and range?

The German government introduced a subsidy for electric vehicles in 2016. The goal was to increase diffusion to have 1 million electric cars on the streets by 2020 and 6 million by 2030. In this section, I quantify the impact of the introduction of this subsidy on the prices and range levels of battery electric vehicles. To do so, I re-compute the market equilibrium in 2018 without the subsidy. As in the case of a shock to the marginal cost of providing range, it is unclear how firms adjust the price and range of their products.[15] There also exist reasons to think that the response to a subsidy may be different from the reaction to a shock to the marginal cost of providing range: A subsidy is equivalent to a decrease in the total marginal cost of producing a product and not specific to particular adjustable product characteristics.

The results are in Table 1.3 and show outcomes from three counterfactuals: Column 3 shows outcomes when both price and range are allowed to adjust. Columns 4 and 5 show outcomes where only price and only range are allowed to adjust, respectively.

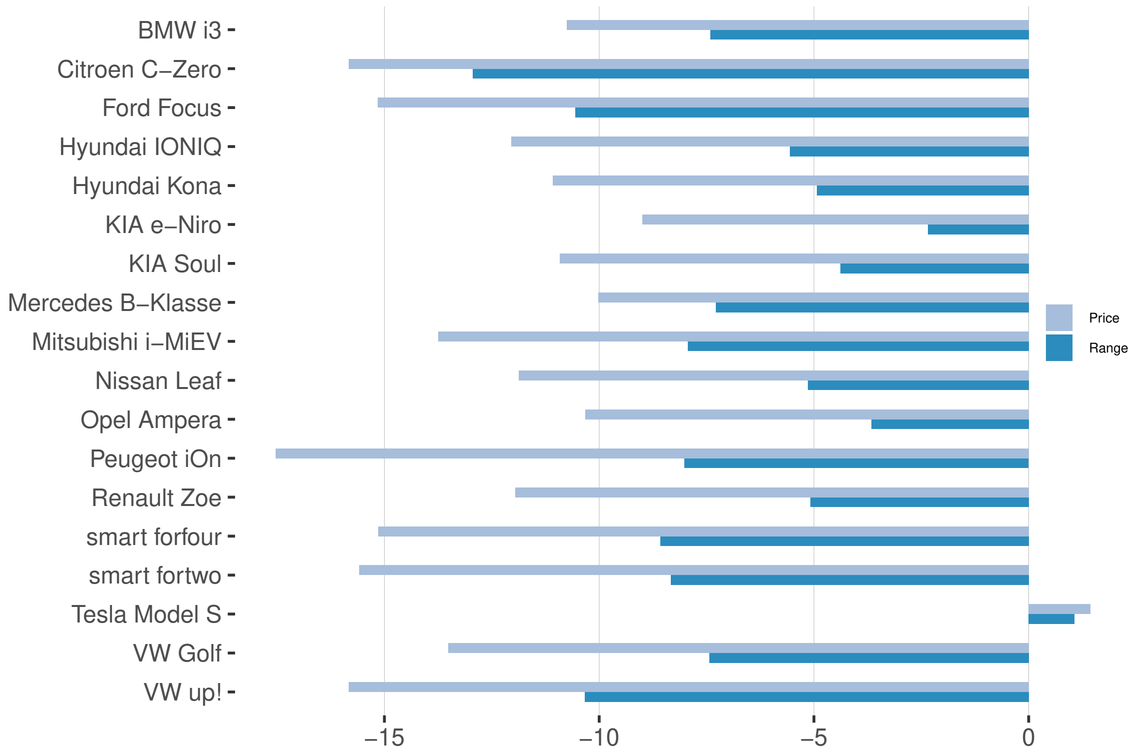
Table 1.3: Difference in market outcomes without the subsidy

	With subsidy	**Without subsidy**		
	Base	Price, range adjust	Only price adjusts	Only range adjusts
Price	34,671	+4,099	+2,268	0
		(+3,806, +5,799)	(+2,219, +2,316)	
Range	259	+15	0	-13
		(+9, +30)		(-21, -10)
MC	26,023	+1,522	0	-1,294
		(+1,235, +2,925)		(-1,371, -1,195)
Markup	10,536	+688	+379	-595
		(+669, +993)	(+331, +427)	(-693, -518)
Sales	34,761	-7,431	-6,707	-4,333
		(-8,222, -6,571)	(-7,422, -5,972)	(-5,174, -3,547)

Note: The table gives the differences from observed outcomes with 95% CIs in parentheses.

[15] In Appendix E.2, I show how the system of first-order conditions can be used to predict the direction of the effects without having to perform a counterfactual analysis.

Figure 1.7: Percentage changes of price and range due to introduction of subsidy



We can see in column 3 that the predictions drawn from the comparative statics are validated. Without the subsidy, both prices and range levels would have been higher on average. The results suggest that on average, price pass-through was more than 100% (the subsidy was €2,000 in 2018). The pass-through rate to price was over 200%. Firms compensated for this over-shifting by lowering the range. Markups also fell in response to the subsidy, meaning that firms sold cheaper cars with less range on which they collected a smaller markup. Column 4 suggests over-shifting of the subsidy in a counterfactual scenario in which firms were only allowed to adjust prices. The pass-through rate is approximately 113% in this case. This rate is slightly higher than that found by Muehlegger and Rapson (2020) for subsidies in California, where pass-through was indistinguishable from 100%. Column 5 indicates that if firms had only been able to change range, the subsidy would have led to an increase in range at a higher markup. In the last line of Table 1.3, we can see that the subsidy increased sales by 7,431 units or approximately 21%. We also see that not accounting for range adjustment leads to an under-prediction of the effect that the subsidy had on sales. In either case, we can conclude that the subsidy is far from generating the diffusion needed to achieve the goal of having 1 million electric vehicles on the streets by 2020. Figure 1.7 shows the product-level effects of the subsidy. We can see that firms decreased the price and range of all subsidized products, and the lone non-subsidized BEV, the Tesla Model S, saw an increase in both price and range in response to the subsidy.

Discussion

The analysis of marginal cost shocks and the subsidy makes apparent two countervailing forces in the market for battery electric vehicles: On the one hand, subsidies put downward pressure on prices, range, and markups. On the other hand, a lower marginal cost of providing range puts upward pressure on prices, range, and markups. The fact that the subsidy leads firms to sell cheaper cars with lower range and that a lower marginal cost of providing such range leads firms to sell more expensive cars with higher range suggests that substitution patterns from combustion cars, as well as the outside option, are different in the two cases. Knowing what substitution patterns look like and how they change in different scenarios is essential for subsidy design.

The countervailing effects of subsidies and a lower marginal cost of range can also help to explain the evolution of price and range over the sample period (see, e.g., Figure 1.2): Until the end of 2015, there was no subsidy in place, meaning that there was only upward pressure on prices and range levels, explaining the increasing slopes of both curves. Starting with the introduction of the subsidy in 2016, we see that prices first plateaued and then decreased, suggesting that the subsidy's negative price effect dominated the positive price effect of the marginal cost drop. The net impact on range stayed positive, even though the increase seems to have slowed down. Overall, the total effect depends on the amount of the subsidy, the magnitude of the marginal cost decrease, consumer preferences, and the current price and range levels. Based on these factors, the total effect of subsidies and changes in the marginal cost of range on price and range may be either positive or negative.

These results raise questions for policymakers regarding subsidy design. The findings suggest that a possibly unintended side effect of flat purchase subsidies is a lower range. On the one hand, using the support scheme to offer lower-range, lower-price products may be desirable for very price-sensitive consumers and allow firms to increase sales. The results from these two counterfactual exercises also raise the question of how a policymaker can achieve different objectives through subsidy design. At what level would consumer surplus be maximized? The ultimate goal of policymakers is to eliminate CO2 emissions from new cars sold to de-carbonize the transport sector. What subsidy scheme achieves minimal CO2 emissions from new sales? Finally, many governments have introduced sales targets for electric cars that they try to meet by maximizing diffusion. A diffusion-maximizing strategy may also make sense when policymakers have dynamic incentives, such as moving down a learning curve. In that case, it can be optimal to forgo emission savings now because moving down the learning curve quickly leads to higher diffusion and higher emission savings in the future. The next section investigates which subsidy schemes achieve different objectives and whether they are compatible with one another.

Incentive-based subsidies

Policymakers in different countries use different subsidy schemes. For instance, the total subsidy in California and China is a function of the driving range or the size of the battery pack (Rokadiya and Yang, 2019). In Germany, on the other hand, the amount does not differ across different EVs. In this section, I compare range-based schemes with schemes that are invariant across different models. In particular, I evaluate subsidies of the form $\lambda_j = \lambda_0 + \lambda_1 r_j$. Note that while simple, this scheme nests both the case of a flat subsidy and a decrease in the marginal cost of providing range. When λ_1 is zero, we recover a simple flat subsidy of the form implemented in Germany. When λ_0 is zero, the subsidy depends purely on the range. In that case, the subsidy is equivalent to a decrease in the marginal cost of providing range. On the other hand, a flat subsidy is equivalent to a general marginal cost decrease. In other words, a flat subsidy lets firms choose how to "interpret" the marginal cost decrease: They can treat it as making range provision cheaper or as reducing the total marginal cost of producing the product. By contrast, a pure range-based subsidy forces firms to treat the subsidy as a decrease in the marginal cost of providing range. One can interpret the intermediate cases where both λ_0 and λ_1 are non-zero as putting weights on a general and a range-specific marginal cost decrease.

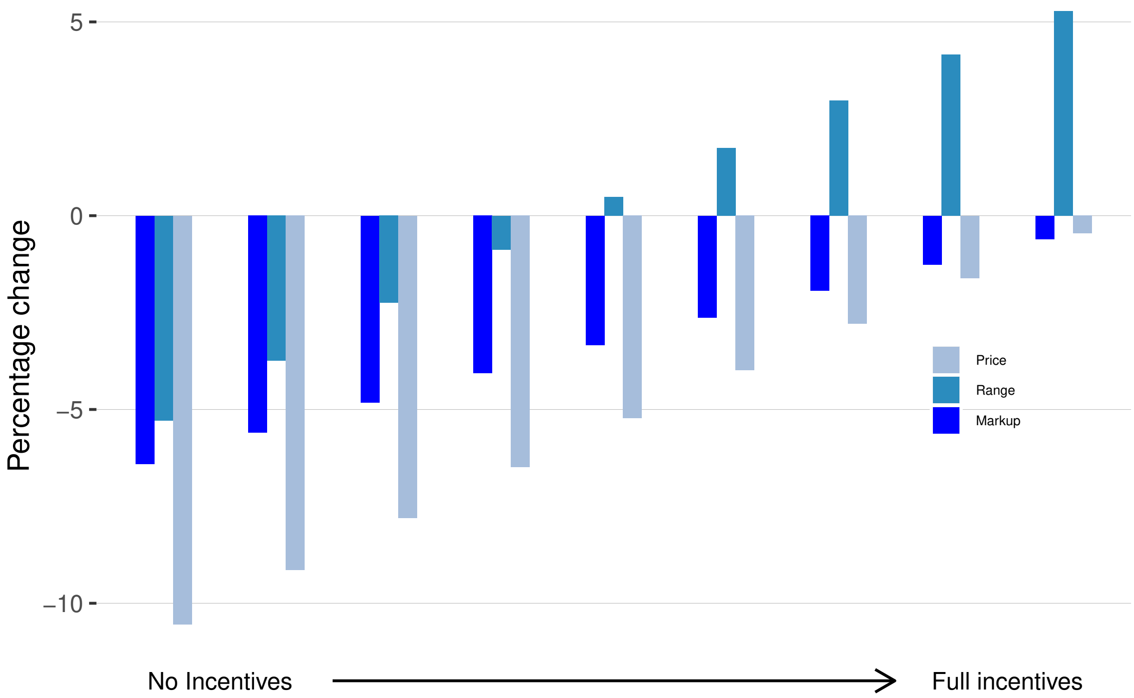
To find the budget-equivalent values for λ_0 and λ_1, I use the following procedure: At a given budget B, I search for values of λ_0, λ_1 that satisfy the budget constraint. I employ a grid search where at each candidate value $(\tilde{\lambda_0}, \tilde{\lambda_1})$, I solve for the counterfactual equilibrium vector of prices and ranges and compute the total cost of the scheme. If the cost is either above or below B, I discard the candidate value, and if the cost is equal to B (up to a small tolerance), I keep it. I perform this search for different values of B: In the first search, B is equal to the observed subsidy scheme's cost in 2018 and subsequently increases in further searches. For each candidate point, I compute the mean price and range of BEVs, the quantity sold of BEVs, consumer surplus[16], and fleet emissions, computed as the sales-weighted CO2 emissions from new cars sold.[17] Note that in the computation of fleet emissions, I assume that BEVs' CO2 emissions are equal to zero. Of course, this assumption is only true if they run exclusively on electricity generated from renewable sources. The assumption is unrealistic in a country such as Germany, where an important part of electricity generation comes from CO2-intensive coal-fired plants. However, there are three reasons why this approach is justified. The first is that it serves as a useful benchmark since it measures the maximum amount by which fleet emissions can decrease. The second is that the main reason why policymakers see electric vehicles as a key instrument in making the transport sector emission-free is that electricity generation itself

[16] Consumer surplus is computed using the log-sum formula: $CS_t = \sum_m \phi_{mt} \sum_i w_i \frac{log(1+\sum_j \exp(\delta_{jmt}+\mu_{ijmt}))}{\alpha_i}$.

[17] I compute fleet emissions as $\sum_j \mathrm{CO2}_j\, q_j$, with $\mathrm{CO2}_j$ being the CO2 emissions of car j, measured in g/km and q_j being the quantity sold of car j.

Figure 1.8: Firm strategies across subsidy schemes

Mean percentage changes at lowest budget



is being de-carbonized. De-carbonized electricity generation means that BEVs will eventually be emission-free, making it a useful benchmark to think of them as zero-emission vehicles. The third reason is that assuming non-zero CO2 emissions from BEVs requires ad hoc assumptions on the electricity mix used and driving behaviour.

I focus on three outcomes in this section: First, I look at CO2 emissions from new car sales. Focusing on this target makes sense, as the ultimate goal of subsidizing BEVs is to de-carbonize the transport sector. The fewer vehicles emitting CO2 sold, the lower are the CO2 emissions from the existing vehicle stock. Second, I focus on diffusion. This target makes sense for two reasons. First, many governments have introduced explicit sales targets for electric vehicles. A diffusion-maximizing approach ensures the achievement of these sales targets. Second, a strategy focusing on maximizing diffusion can also be a static approximation to a dynamic optimization problem: A policymaker quickly wants to move down a learning curve. A diffusion-maximizing strategy can approximate well the desire to move down the learning curve swiftly in the early phase of adoption. An interpretation of sales targets can be that the policymaker simplifies the complicated dynamic optimization problem by defining short- and medium-run sales targets that allow the industry to move down the learning curve quickly. Finally, I take into account distributional aspects by looking at consumer surplus along income deciles.

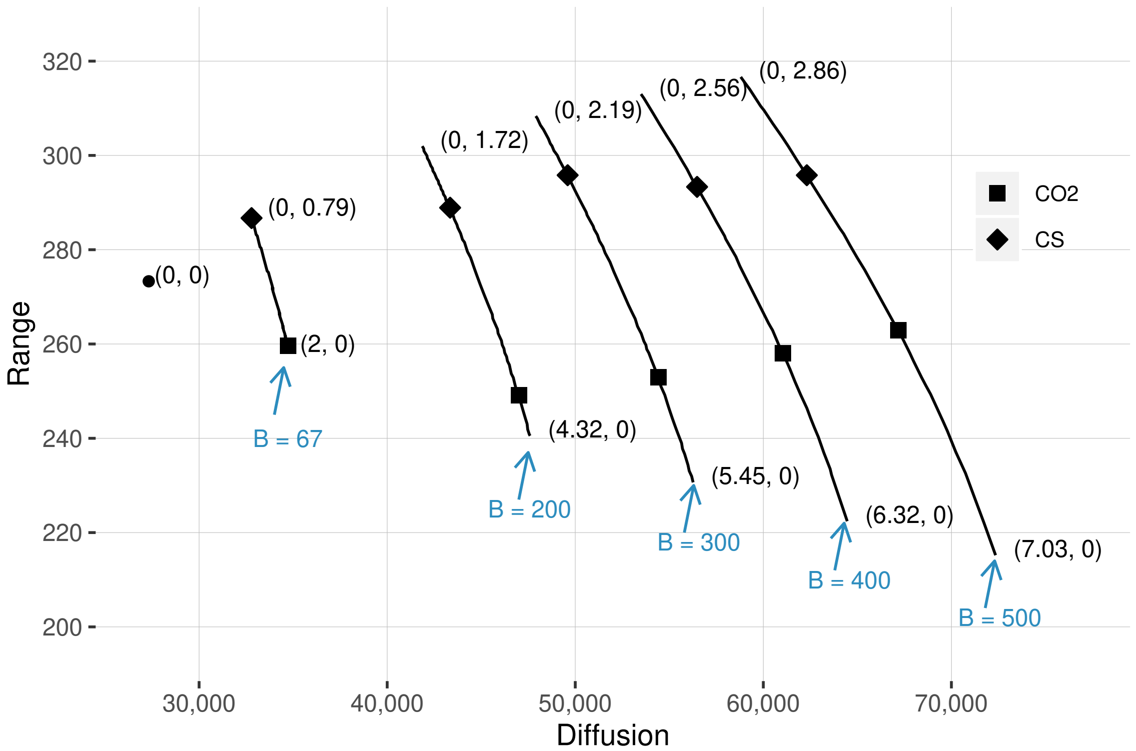
Firms adopt two strategies, depending on the subsidy's design: First, when the subsidy is flat, we recover the previous section's result: Firms decrease both price and range and collect a lower markup on their BEVs. Firms also employ this strategy when facing schemes with low-powered incentives on range. Second, when subsidies put more incentives on range, firms use a mix of lower prices and higher range. The decrease in markups is smaller, meaning that at schemes with stronger incentives, firms sell relatively more expensive BEVs with a higher range and a relatively higher markup. We do not recover the result from Section 1.6 that firms increase both price and range and earn a higher markup on their BEVs. Note that in Section 1.6, a (flat) subsidy is already in place. The decrease in the marginal cost is then akin to replacing the flat subsidy with a mixed scheme. Increasing incentives always increases range and markups and decreases prices. Figure 1.8 shows how firms react to different subsidy schemes when the budget is €67 million. While the magnitudes of the effects change at higher budgets, the pattern described above does not: Flat subsidies and low-incentive mixed schemes always lead to cheaper BEVs with a lower range on which firms collect a lower markup. Pure range-based and high-incentive mixed schemes always lead to relatively more expensive BEVs with a higher range on which firms collect a relatively higher markup. It is crucial to understand firms' strategic reactions to different subsidy schemes because these strategies lead to different BEVs on the market and different substitution patterns. These substitution patterns shape policy objectives such as emissions from new car sales and diffusion.

Policy outcomes

Figure 1.9 shows policy outcomes under different subsidy schemes. The plot shows iso-budget curves for different budget levels and the "initial point" without subsidies from Table 1.3. The black diamonds denote the subsidy scheme maximizing consumer surplus at a given budget, whereas the black squares represent the scheme minimizing fleet emissions at a given budget. We can see that the flat subsidy *always* maximizes diffusion, suggesting that lower-range, cheap BEVs induce the most substitution from other cars and the outside option. Also, we see that fleet emissions are minimized at intermediate schemes. Maximizing diffusion of zero-emission BEVs and minimizing fleet emissions are not equivalent. Minimizing emissions calls for different substitution patterns compared to those generated by a flat subsidy. More substitution from polluting cars makes up for the lower overall substitution. A mixed subsidy scheme placing some incentives on range provision achieves this substitution from more polluting cars. Finally, schemes with strong incentives on range always maximize consumer surplus. Consumers care more about range than about price in relative terms to prefer a market outcome at which BEVs have a high range at higher prices, rather than a market outcome where the opposite holds. Together, these findings show that maximizing diffusion, minimizing fleet emissions, and maximizing consumer surplus are mutually exclusive at any budget. The policymaker can achieve these two goals separately or accomplish a mix of these goals, as any subsidy scheme

Figure 1.9: Policy outcomes at different subsidy schemes (λ_0, λ_1) and budgets

Subsidy in €1,000, budget in €million



always increases consumer surplus, decreases fleet emissions, and increases diffusion.

Table 1.4: Substitution patterns by engine type

Budget:	**€64 million**			**€200 million**			**€300 million**			**€400 million**			**€500 million**		
	Diffusion	Range	CO2	Diffusion	Range	CO2	Diffusion	Range	CO2	Diffusion	Range	CO2	Diffusion	Range	CO2
Percentage															
Outside option	77.26	73.34	77.15	79.04	73.99	78.34	80.08	74.37	78.43	80.96	74.71	78.46	80.96	74.71	78.46
ICE	17.93	21.05	18.02	16.71	20.64	17.25	15.99	20.4	17.26	15.36	20.2	17.29	15.36	20.2	17.29
PHEV	4.8	5.61	4.83	4.25	5.37	4.41	3.93	5.22	4.31	3.67	5.09	4.24	3.67	5.09	4.24
Absolute															
Outside option	5,904	4,091	5,869	16,187	10,908	15,575	23,340	15,456	21,418	30,234	19,705	26,620	30,234	19,705	24,351
ICE	1,370	1,174	1,371	3,423	3,043	3,430	4,660	4,240	4,714	5,737	5,328	5,868	5,737	5,328	5,779
PHEV	367	313	367	870	792	876	1,147	1,085	1,177	1,372	1,343	1,439	1,372	1,343	1,433
Scheme															
λ_0	2,000	0	1,950	4,315	0	3,900	5,450	0	4,400	6,320	0	4,650	7,028	0	4,800
λ_1	0	750	25	0	1,720	210	0	2,188	542	0	2,555	866	0	2859	1,156

Percentage terms may not equal 100% due to rounding errors.

Table 1.4 shows substitution patterns from the outside option, combustion cars (ICEs), and PHEVs towards BEVs under the schemes maximizing diffusion and range, respectively, and minimizing fleet emissions (CO2) at the different budgets. We can see several patterns that help explain Figure 1.9: First, substitution from the outside option is always highest at the flat scheme and lowest at the pure incentive-based scheme. Second, substitution from the inside goods in relative terms is highest at the pure incentive-based scheme and is lowest at the flat scheme. Lower substitution from inside goods is more than made up for by higher substitution from the outside good at the flat scheme, thus maximizing diffusion. It is a profit-maximizing strategy for firms to use a flat subsidy to sell cheaper products at a lower range. This strategy has a market-expanding effect and mainly attracts consumers who chose not to purchase a new car before. On the other hand, pure incentive-based subsidies make it profit-maximizing for firms to sell more expensive cars with a higher range. This strategy attracts consumers who value the higher range provided and previously purchased a non-BEV inside good. Firms use a similar strategy at high-incentive mixed schemes that maximize consumer surplus, suggesting that consumers, on average, have a relatively higher sensitivity to range than to price. Finally, Table 1.4 also explains why intermediate schemes (or the flat scheme for the lowest budget) minimize fleet emissions: Combustion cars have higher CO2 emissions than PHEVs. Together with the assumption that consumers who choose the outside option cause zero CO2 emissions[18], this means that fleet emissions will be lowest at the point where substitution from combustion cars is the highest. Maximal substitution from combustion cars occurs at mixed schemes.

Table 1.5 shows the market outcomes of the flat schemes and the pure incentive-based schemes at different budgets. When the budget is set to €200 million, for instance, the flat subsidy ($\lambda_1 = 0$) leads to a 33km, 12%, decrease in range and a 76% increase in diffusion, or 20,480 cars in absolute terms. A pure incentive-based subsidy ($\lambda_0 = 0$), on the other hand, leads to a 29km, or 11%, increase in range but only a 55% increase in diffusion, or 14,742 cars in absolute terms. In other words, an extra 5,738 units "costs" the policymaker 62km of range, equal to around 23% of the range absent any subsidies. We can also see that the flat subsidy raises total firm profits by more than the pure incentive-based subsidy. The flat subsidy gives firms more flexibility in "interpreting" the subsidy: It is equivalent to a general decrease in the marginal cost of producing a BEV. In contrast, the pure incentive-based subsidy is equivalent to a decrease in the range-specific part of marginal cost.

Distributional aspects

Table 1.5 makes clear that the market outcomes for BEVs are significantly different between the flat- and the pure range-based subsidy schemes. Flat schemes produce cheaper BEVs with a lower range, whereas incentive-based schemes create more expensive BEVs with a higher range. These two outcomes likely mean that different consumers buy BEVs in the two cases. The demand model results yield a lower price sensitivity from consumers with higher income, meaning that their willingness to pay for range is higher. This greater willingness to pay at higher income deciles suggests that consumer surplus alone hides important distributional

[18] This assumes that the consumers attracted from the outside option did not use modes of transport emitting CO2 emissions before.

Table 1.5: Market outcomes at different budgets, flat versus range-based subsidies

Budget:		**€ 64 million**		**€ 200 million**		**€ 300 million**		**€ 400 million**		**€ 500 million**	
	Base	Incentive	Flat	Incentive	Flat	Incentive	Flat	Incentive	Flat	Incentive	Flat
Sales	27,119	+5,578	+7,642	+14,742	+20,480	+20,782	+29,147	+26,377	+37,343	+31,680	+45,230
Range	273.16	+14	-14	+29	-33	+35	-43	+40	-51	+43	-58
Price	38,756	-178	-4,085	-816	-8,805	-1,286	-11,152	-1,721	-12,988	-2,123	-14,500
MC	27,538	+1,866	-1,515	+3,900	-3,254	+4,864	-4,125	+5,592	-4,816	+6,175	-5,392
Markup	11,218	-38	-682	-153	-1,476	-236	-1,879	-313	-2,201	-384	-2,470
Profits	30,792	+79	+81	+139	+147	+178	+189	+213	+227	+246	+262
λ_0	0	+790	+0	+1,720	+0	+2,188	+0	+2,555	+0	+2,858	+0
λ_1	0	0	2,000	0	4,315	0	5,450	0	6,323	0	7,028

Table 1.6: Preferences for subsidy scheme across income deciles

Budget:	**€ 64 million**	**€ 200 million**	**€ 300 million**	**€ 400 million**	**€ 500 million**
Scheme					
Deciles preferring					
Flat	9	9	8	8	8
Mixed	0	0	1	1	1
Incentive	1	1	1	1	1
Marginal decile					
Flat to Mixed	-	-	8-9	8-9	8-9
Mixed to Incentive	-	-	9-10	9-10	9-10
Flat to Incentive	9-10	9-10	-	-	-

effects. Table 1.6 illustrates these distributional effects. The table reports the number of income deciles preferring either a flat, a mixed, or a pure incentive-based scheme along with the "marginal deciles" between which preferences for different schemes switch. We can see that most income deciles prefer flat schemes, with only the top decile having a preference for pure incentive-based schemes. However, the gain in consumer surplus in this decile is enough to make overall consumer surplus maximal at schemes that put strong incentives on range. We can also see that preferences are very polarized in that only the 9th income decile has a preference for a mixed scheme (at a budgets over € 200 million). In contrast, all other income deciles either prefer a flat or a true incentive-based scheme. These findings suggest that there are important distributional consequences to consider when designing subsidies. A policymaker can target different consumer segments with different schemes.

Discussion

These results suggest that at the observed levels of price, range and marginal cost in 2018, a policymaker interested in maximizing diffusion merely needs to introduce flat subsidy schemes. On the other hand, a policymaker interested in minimizing fleet emissions would take another strategy, employing a mixed scheme that maximizes substitution from more polluting combustion cars. A high-incentive mixed subsidy maximizes consumer surplus, even though this hides important distributional aspects. In summary, policymakers face a trade-off between diffusion, emissions, but can address distributional aspects. This suggests that a policymaker needs to be mindful when deciding on a policy objective as it may have unintended consequences on other outcomes. The results also suggest, however, that a policymaker can always achieve a mix of the three goals as a subsidy *always* increases consumer surplus for all consumers, *always* increases diffusion, and *always* reduces fleet emissions. What holds regardless of the level of prices, ranges, and marginal cost of range provision is that a policymaker intending to maximize diffusion with a subsidy should be mindful of the impact of price and range on firms' intensive and extensive margins. A subsidy can lead a firm to use three strategies, two of which we have seen in this section: First, a firm could use the subsidy to decrease both price and range. Second, the firm could use the subsidy to decrease price and increase range. Third, a firm could use the subsidy to increase both the price and range. These three strategies will generate different substitution patterns from polluting cars and the outside option, leading to different market outcomes.

The downward trend in the marginal cost of providing range puts upward pressure on price and range. As long as range provision continues to become cheaper, a flat or low-incentive subsidy's negative effect on the range is likely to be mitigated or even dominated by this range-enhancing effect.

For the German market studied here, the results suggest that flat subsidies were indeed diffusion maximizing, albeit with a moderate overall effect. The findings also indicate that increasing diffusion to a level that would bring the electric vehicle stock close to 1 million by 2020 necessitates a substantial increase in the subsidy amount.

This section's findings are also relevant to other markets: Policymakers often subsidize access to necessary infrastructures such as water and electricity in developing countries. These subsidies may have adverse effects on quality (McRae, 2015). Another example is newspaper markets, where quality may be affected by subsidies aimed at increasing readership numbers (Battaggion and Vaglio, 2018). In these cases, policymakers need to be mindful of relative preferences for price and quality, the cost of quality provision, and the impact of these factors on firms' intensive and extensive margins. Ultimately, a subsidy's effect on price and an adjustable product attribute is an empirical question that calls for a case-by-case evaluation. In general, this section shows that subsidy design in a multi-product oligopoly when firm reactions can be multi-dimensional can lead to very different strategic reactions by firms, affecting substitution patterns that ultimately shape market outcomes.

1.7 Conclusion

In this paper, I study how firms adjust the price and range of electric vehicles in response to subsidies and changes in marginal cost. Falling input prices and subsidies characterize the electric vehicle market. Even though understanding how input prices and subsidies are passed through to price and range and how they affect the diffusion of electric vehicles is essential for proper subsidy design, there is little evidence on pass-through when price and range are endogenous.

I develop a structural model of demand and supply for new cars, and estimate it using a novel data set on state-level new car sales in Germany. On the demand side, consumers choose between cars of different engine types. The demand side allows for rich substitution patterns across electric and combustion cars. On the supply side, firms compete in prices and can set the range of their battery electric vehicles (BEVs). The model provides a framework for analyzing the impact of subsidies in imperfectly competitive markets when firms choose the price and product attributes.

I find that the marginal cost of providing range has decreased by approximately 33% over the sample period. I use the estimated model to analyze how firms adjust price and range in response to cheaper range provision and subsidies. The lower marginal cost of range provision increased the price and range of BEVs, with firms collecting a higher markup. Conversely, a flat subsidy introduced in Germany led to cheaper BEVs with a lower range on which firms collect a lower markup. The subsidy increased sales by approximately 27% in 2018, far from sufficient to meet the governments' sales targets.

I then compare the flat subsidy imposed in Germany to alternative schemes used in other countries and their effect on policy goals. I find that policymakers face a trade-off between maximizing diffusion, minimizing emissions, but can address distributional concerns. Different substitution patterns at different schemes drive this result. These substitution patterns ultimately determine market outcomes. On the one hand, flat subsidies and schemes with low-incentive mixed schemes induce firms to employ a strategy of selling BEVs with less range at a lower price, capturing a large number of consumers on which firms collect a smaller markup. On the other hand, pure range-based subsidies and high-incentive mixed schemes induce firms to sell BEVs with more range at a relatively higher price, attracting consumers with a high willingness to pay on which firms collect a relatively higher markup. The findings suggest that a policymaker can always achieve a mix of the three policy goals as the subsidies always increase consumer surplus and diffusion and always decrease fleet emissions.

This result also has direct implications for distributional effects: As consumers with a higher income have a higher willingness to pay for range, they are better off under high-incentive schemes. Consumers in lower-income deciles, to the contrary, are better off under a flat subsidy scheme. On the one hand, there may be important distributional effects from different subsidy schemes. On the other hand, targeting specific consumer groups can be compatible with a given policy objective. The diffusion-maximizing subsidy scheme coincides with the scheme maximizing consumer surplus for lower-income groups, for instance.

The results have implications for policymakers. It is crucial to understand substitution patterns generated by different subsidy schemes. Consumer preferences ultimately drive these substitution patterns for range and price and the marginal cost of providing range. These insights generalize to other markets in which firms can adjust one or more product attributes in response to subsidies.

My paper leaves scope for future work. First, I do not directly explore dynamic incentives that may exist due to learning effects. Second, I take the product portfolio of firms as given. Recent years have seen the introduction of a large number of new EV models. Endogenizing the product portfolio may be necessary to understand how firms react to subsidies and cost changes by (not) introducing new products. Finally, firms have been increasingly offering models with different range specifications. Firms offering menus of price and range add an angle to range provision as firms may distort price and range within their menu.

References

Maria Rosa Battaggion and Alessandro Vaglio. Newspapers and public grants: A matter of quality. *Scottish Journal of Political Economy*, 65(1):27–38, 2018.

Arie Beresteanu and Shanjun Li. Gasoline prices, government support, and the demand for hybrid vehicles in the united states. *International Economic Review*, 52(1):161–182, 2011.

Steven Berry and Philip Haile. Identification in differentiated products markets using market level data. *Econometrica*, 82(5):1749–1797, 2014.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.

Daniel Brunner, Florian Heiss, André Romahn, and Constantin Weiser. *Reliable estimation of random coefficient logit demand models.* DICE Discussion Papers 267, 2017.

Jeremy I Bulow and Paul Pfleiderer. A note on the effect of cost changes on prices. *Journal of Political Economy*, 91(1):182–185, 1983.

Sanya Carley, Rachel M Krause, Bradley W Lane, and John D Graham. Intent to purchase a plug-in electric vehicle: A survey of early impressions in large us cites. *Transportation Research Part D: Transport and Environment*, 18:39–45, 2013.

Gary Chamberlain. Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of Econometrics*, 34(3):305–334, 1987.

Christopher Conlon and Jeff Gortmaker. Best practices for differentiated products demand estimation with pyblp. *The RAND Journal of Economics*, Forthcoming.

Gregory Crawford, Oleksandr Shcherbakov, and Matthew Shum. Quality overprovision in cable television markets. *American Economic Review*, 109(3):956–95, 2019.

Xavier D'Haultfœuille, Isis Durrmeyer, and Philippe Février. Automobile prices in market equilibrium with unobserved price discrimination. *The Review of Economic Studies*, 86(5): 1973–1998, 2019.

Jean-Pierre Dubé, Jeremy T Fox, and Che-Lin Su. Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation. *Econometrica*, 80(5):2231–2267, 2012.

Jean-Pierre H Dubé, Ali Hortaçsu, and Joonhwi Joo. Random-coefficients logit demand estimation with zero-valued market shares. NBER Working Papers 26795, National Bureau of Economic Research, 2020.

Isis Durrmeyer and Mario Samano. To rebate or not to rebate: Fuel economy standards versus feebates. *The Economic Journal*, 128(616):3076–3116, 2018.

Ying Fan. Ownership consolidation and product characteristics: A study of the us daily newspaper market. *American Economic Review*, 103(5):1598–1628, 2013.

Amit Gandhi and Jean-François Houde. Measuring Substitution Patterns in Differentiated Products Industries. NBER Working Papers 26375, National Bureau of Economic Research, 2019.

Amit Gandhi, Zhentong Lu, and Xiaoxia Shi. Estimating demand for differentiated products with error in market shares. 2013.

William H Green, Robert C Armstrong, Moshe Ben-Akiva, John Heywood, Christopher Knittel, Sergey Paltsev, Brian Reimer, Chintan Vaishnav, Jinhua Zhao, Eytan Gross, et al. Insights into future mobility: A report from the mobility of the future study. 2019.

Laura Grigolon, Mathias Reynaert, and Frank Verboven. Consumer valuation of fuel costs and tax policy: Evidence from the european car market. *American Economic Journal: Economic Policy*, 10(3):193–225, 2018.

Stephane Hess, Kenneth E Train, and John W Polak. On the use of a modified latin hypercube sampling (mlhs) method in the estimation of a mixed logit model for vehicle choice. *Transportation Research Part B: Methodological*, 40(2):147–163, 2006.

Stephen P. Holland, Erin T. Mansur, and Andrew J. Yates. The Electric Vehicle Transition and the Economics of Banning Gasoline Vehicles. NBER Working Papers 26804, National Bureau of Economic Research, February 2020.

I-Yun Lisa Hsieh, Menghsuan Sam Pan, Yet-Ming Chiang, and William H Green. Learning only buys you so much: practical limits on battery price reduction. *Applied Energy*, 239: 218–224, 2019.

Donghun Kim and Ronald W Cotterill. Cost pass-through in differentiated product markets: The case of us processed cheese. *The Journal of Industrial Economics*, 56(1):32–48, 2008.

Thomas Klier and Joshua Linn. New-vehicle characteristics and the cost of the corporate average fuel economy standard. *The RAND Journal of Economics*, 43(1):186–213, 2012.

Christopher R Knittel. Automobiles on steroids: Product attribute trade-offs and technological progress in the automobile sector. *American Economic Review*, 101(7):3368–99, 2011.

Christopher R Knittel and Konstantinos Metaxoglou. Estimation of random-coefficient demand models: two empiricists' perspective. *Review of Economics and Statistics*, 96(1):34–59, 2014.

Benjamin Leard, Joshua Linn, and Katalin Springel. Pass-through and welfare effects of regulations that affect product attributes. *Working Paper*, 2019.

Jing Li. Compatibility and investment in the us electric vehicle market. *Working Paper*, 2019.

Shanjun Li, Joshua Linn, and Erich Muehlegger. Gasoline Taxes and Consumer Behavior. *American Economic Journal: Economic Policy*, 6(4):302–342, 2014.

Shanjun Li, Lang Tong, Jianwei Xing, and Yiyi Zhou. The market for electric vehicles: indirect network effects and policy design. *Journal of the Association of Environmental and Resource Economists*, 4(1):89–133, 2017.

Zhentong Lu, Xiaoxia Shi, and Amit Gandhi. Estimating demand for differentiated products with zeroes in market share data. *Working Paper*, 2019.

Nic Lutsey and Michael Nicholas. Update on electric vehicle costs in the united states through 2030. *International Council on Clean Transportation (ICCT)*, 2019.

Eric Maskin and John Riley. Monopoly with incomplete information. *The RAND Journal of Economics*, 15(2):171–196, 1984.

Shaun McRae. Infrastructure quality and the subsidy trap. *American Economic Review*, 105 (1):35–66, 2015.

Erich Muehlegger and David S. Rapson. Subsidizing Mass Adoption of Electric Vehicles: Quasi-Experimental Evidence from California. NBER Working Papers 25359, National Bureau of Economic Research, 2020.

Michael Mussa and Sherwin Rosen. Monopoly and product quality. *Journal of Economic Theory*, 18(2):301–317, 1978.

Aviv Nevo. Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2): 307–342, 2001.

Björn Nykvist and Måns Nilsson. Rapidly falling costs of battery packs for electric vehicles. *Nature climate change*, 5(4):329–332, 2015.

Giulia Pavan. Green Car Adoption and the Supply of Alternative Fuels. TSE Working Papers 17-875, Toulouse School of Economics (TSE), 2017.

Mathias Reynaert. Abatement strategies and the cost of environmental regulation: Emission standards on the european car market. *The Review of Economic Studies*, Forthcoming.

Mathias Reynaert and James Sallee. Who benefits when firms game corrective policies? *American Economic Journal: Economic Policy*, Forthcoming.

Mathias Reynaert and Frank Verboven. Improving the performance of random coefficients demand models: the role of optimal instruments. *Journal of Econometrics*, 179(1):83–98, 2014.

Zeinab Rezvani, Johan Jansson, and Jan Bodin. Advances in consumer electric vehicle adoption research: A review and research agenda. *Transportation research part D: transport and environment*, 34:122–136, 2015.

S Rokadiya and Z Yang. Overview of global zero-emission vehicle mandate programs. *International Council on Clean Transportation (ICCT)*, 2019.

Brandon Schoettle and Michael Sivak. Resale values of electric and conventional vehicles: Recent trends and influence on the decision to purchase a new vehicle. 2018. "`http://umich.edu/~umtriswt/PDF/SWT-2018-4.pdf`".

Eytan Sheshinski. Price, quality and quantity regulation in monopoly situations. *Economica*, 43(170):127–137, 1976.

Michael Spence. Monopoly, quality, and regulation. *The Bell Journal of Economics*, pages 417–429, 1975.

Katalin Springel. Network externality and subsidy structure in two-sided markets: Evidence from electric vehicle incentives. *Working Paper*, 2020.

Marc Steen, Natalia Lebedeva, Franco Di Persio, and L Boon-Brett. Eu competitiveness in advanced li-ion batteries for e-mobility and stationary storage applications–opportunities and actions. *Publ. Off. Eur. Union*, 44, 2017.

Nicholas Stern. The effects of taxation, price control and government contracts in oligopoly and monopolistic competition. *Journal of Public Economics*, 32(2):133–158, 1987.

Jeff Thurk. Sincerest form of flattery? product innovation and imitation in the european automobile industry. *The Journal of Industrial Economics*, 66(4):816–865, 2018.

E Glen Weyl and Michal Fabinger. Pass-through as an economic tool: Principles of incidence under imperfect competition. *Journal of Political Economy*, 121(3):528–583, 2013.

Jianwei Xing, Benjamin Leard, and Shanjun Li. What does an electric vehicle replace? NBER Working Papers 25771, National Bureau of Economic Research, 2019.

Zifei Yang, Peter Slowik, Nic Lutsey, and Stephanie Searle. Principles for effective electric vehicle incentive design. *Internatnional Council Clean Transportation*, 2016.

Appendices

Appendix A

Additional Figures and Tables

Table A.1: Summary statistics

Mean values of key characteristics

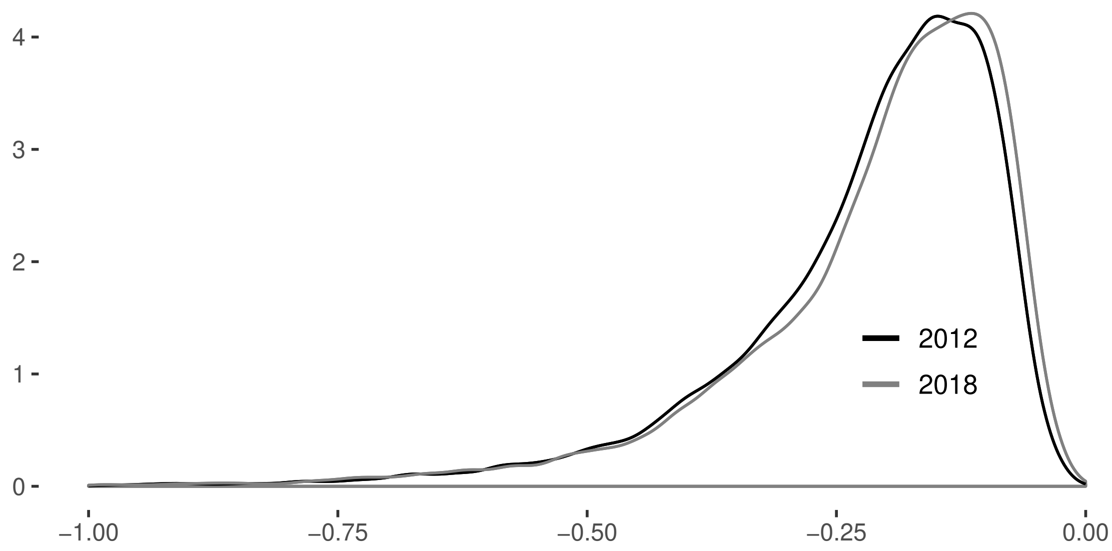
Variable	2012	2013	2014	2015	2016	2017	2018
BEV							
Price	30,490	31,295	35,392	32,569	37,104	37,200	34,671
Range (in km)	168	173	202	196	213	246	259
Fuel Cost	4.02	4.34	4.37	4.19	4.24	4.28	4.21
Acceleration	2.8	2.98	3.19	2.96	3.31	3.26	2.94
Weight	1,581	1,662	1,797	1,797	1,867	1,902	1,841
Footprint	6.01	6.4	6.78	6.78	7.03	7.13	6.97
Doors	4.5	4.7	4.85	4.85	4.86	4.88	4.89
Number of Products	6	10	13	13	14	16	18
Sales	2,100	5,517	9,044	13,234	12,201	25,593	34,629
PHEV							
Price	43,288	48,472	44,265	56,007	57,479	54,651	57,126
Range (in km)	54	53	52	44	40	45	45
Fuel Cost	5.29	5.64	5.76	5.77	5.57	5.58	5.89
Acceleration	4.58	5.16	5.02	5.81	5.82	5.81	5.95
Weight	1,988	2,160	2,143	2,408	2,476	2,425	2,449
Footprint	7.93	8.17	8.04	8.53	8.66	8.66	8.74
Doors	5	5	5	5	4.87	4.86	4.79
Number of Products	2	3	6	11	15	22	24
Sales	1,148	1,079	2,671	8,248	10,614	25,374	25,841
ICE							
Price	32,582	32,873	33,914	33,881	34,653	33,669	33,652
Range (in km)	995	1,018	1,039	1,057	1,063	1,023	997
Fuel Cost	10.06	9.32	8.62	7.6	6.98	7.47	8.01
Acceleration	5.29	5.32	5.41	5.44	5.62	5.76	5.74
Weight	2,023	2,035	2,044	2,043	2,031	2,008	2,017
Footprint	8	8.04	8.07	8.08	8.1	8.09	8.12
Doors	4.43	4.48	4.52	4.55	4.52	4.58	4.63
Number of Products	233	233	227	222	214	213	215
Sales	2,739,581	2,569,876	2,651,415	2,767,185	2,855,922	2,864,409	2,819,762
Stations							
Number of Charging Stations	1,116	1,466	2,243	3,530	6,053	9,803	16,307

Table A.2: First Stage Estimates

	Price		Range		Range x Trend		Stations	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Exogenous Charac.								
Fuel Cost	-0.603	(0.025)	0.000	(0.000)	0.005	(0.002)	0.001	(0.001)
Footprint	4.149	(0.121)	0.065	(0.003)	0.248	(0.018)	-0.001	(0.004)
Acceleration	2.886	(0.044)	-0.024	(0.002)	-0.118	(0.011)	-0.004	(0.002)
Doors	1.279	(0.075)	-0.035	(0.002)	-0.175	(0.010)	-0.002	(0.002)
BEV	-2.058	(1.244)	0.732	(0.073)	-12.668	(0.511)	-0.850	(0.083)
PHEV	0.464	(0.711)	-0.028	(0.014)	-0.553	(0.121)	-0.020	(0.063)
Own State	-0.001	(0.247)	0.000	(0.009)	-0.001	(0.057)	0.057	(0.017)
Trend	-0.722	(0.018)	-0.008	(0.000)	-0.018	(0.003)	0.001	(0.001)
PHEV								
Range x PHEV	-2.459	(0.969)	0.257	(0.070)	-0.871	(0.473)	0.153	(0.154)
Cost shifters								
Station Subsidies	0.005	(0.016)	0.000	(0.001)	0.004	(0.004)	0.019	(0.002)
Differentiation IVs								
Price-quadratic-own	0.332	(0.010)	-0.001	(0.000)	-0.006	(0.002)	-0.001	(0.000)
Price-local-own-nest	8.711	(2.129)	-0.918	(0.097)	-4.624	(0.551)	-0.130	(0.086)
Engine-local-own-nest	-11.343	(1.688)	0.091	(0.025)	0.284	(0.140)	0.041	(0.059)
Engine-quadratic-own-nest	-12.897	(0.949)	0.036	(0.008)	0.287	(0.049)	0.048	(0.022)
Engine-local-rival-nest	-4.973	(0.199)	-0.007	(0.002)	0.022	(0.014)	0.009	(0.005)
Acceleration-quadratic-rival	2.276	(0.116)	0.082	(0.007)	0.424	(0.044)	0.001	(0.009)
Acceleration-quadratic-rival-nest	-2.432	(0.138)	-0.086	(0.007)	-0.445	(0.047)	0.001	(0.010)
Footprint-local-own	-0.198	(0.887)	1.116	(0.059)	5.233	(0.337)	0.033	(0.048)
Engine-local-rival	-1.135	(0.205)	-0.112	(0.008)	-0.684	(0.051)	-0.065	(0.027)
EV efficiency-quadratic-rival-nest	0.238	(0.033)	0.005	(0.001)	0.129	(0.007)	0.049	(0.006)
Fuel efficiency-quadratic-own	0.382	(0.110)	-0.005	(0.001)	-0.040	(0.005)	-0.004	(0.002)
PHEV-count-own	0.012	(0.001)	0.000	(0.000)	-0.002	(0.000)	0.000	(0.000)
PHEV range-local-own	-1.459	(0.112)	0.039	(0.003)	0.598	(0.025)	0.146	(0.018)
BEV count-local-rival	-0.156	(0.058)	0.062	(0.005)	1.366	(0.036)	0.088	(0.005)
Firm FE	X		X		X		X	
Class FE	X		X		X		X	
Body FE	X		X		X		X	
State FE	X		X		X		X	
SW F-Stat	140.664		209.89		96.623		52.472	
Observations	28,288		28,288		28,288		28,288	

Note: This table presents first stage estimates for each of the endogenous charateristics. The Sanderson-Windmeijer multivariate F-test is reported for each endogenous vairable.

Table A.3: Kernel density estimation of price sensitivity



Appendix B

Full demand and supply estimates

Table B.1: Full demand and marginal cost estimates

Utility			Marginal Cost		
	Coefficient	SE		Coefficient	SE
Mean Utility			**Range Provision**		
(Intercept)	-9.705	(0.296)	Intercept	0.813	(0.026)
Range	1.772	(0.223)	Trend	-0.070	(0.006)
Range x Trend	-0.118	(0.024)	**Baseline Marginal Cost**		
Stations	0.610	(0.156)	Intercept	1.534	(0.133)
Fuel Cost	-0.281	(0.027)	Weight	0.274	(0.041)
Footprint	0.504	(0.050)	Fuel Efficiency	-0.041	(0.006)
Acceleration	0.295	(0.025)	KW	0.005	(0.000)
Doors	-0.195	(0.027)	Footprint	0.096	(0.019)
BEV	-8.285	(1.539)	BEV	-0.578	(0.043)
PHEV	-5.901	(1.482)	PHEV	0.189	(0.025)
Own State	1.191	(0.072)	2013	-0.011	(0.013)
Trend	-0.114	(0.010)	2014	-0.026	(0.014)
Audi	2.809	(0.087)	2015	-0.057	(0.014)
BMW	3.173	(0.091)	2016	-0.043	(0.014)
Chevrolet	0.468	(0.125)	2017	-0.026	(0.014)
Citroen	0.272	(0.090)	2018	-0.034	(0.014)
Dacia	0.904	(0.164)	Audi	-0.049	(0.052)
Daihatsu	-0.499	(0.175)	BMW	0.028	(0.054)
Dodge	-3.655	(0.305)	Chevrolet	-0.309	(0.068)

Table B.1: Demand and marginal cost estimates *(continued)*

	Coefficient	SE		Coefficient	SE
Fiat	-0.119	(0.103)	Citroen	-0.153	(0.055)
Ford	1.652	(0.095)	Dacia	-0.843	(0.064)
Honda	1.484	(0.095)	Daihatsu	-0.250	(0.051)
Hyundai	1.536	(0.096)	Dodge	-0.340	(0.093)
Jeep	0.614	(0.107)	Fiat	-0.225	(0.053)
KIA	1.081	(0.092)	Ford	-0.234	(0.055)
Lada	-0.679	(0.188)	Honda	-0.112	(0.056)
Lancia	-1.371	(0.126)	Hyundai	-0.205	(0.054)
Land Rover	1.609	(0.110)	Jeep	-0.151	(0.055)
Mazda	2.170	(0.085)	KIA	-0.242	(0.052)
Mercedes	3.066	(0.096)	Lada	-0.525	(0.063)
MINI	1.797	(0.252)	Lancia	-0.185	(0.054)
Mitsubishi	1.073	(0.107)	Land Rover	-0.143	(0.053)
Nissan	1.143	(0.103)	Mazda	-0.142	(0.053)
Opel	1.596	(0.095)	Mercedes	-0.048	(0.054)
Peugeot	0.822	(0.090)	MINI	-0.011	(0.062)
Renault	1.447	(0.092)	Mitsubishi	-0.199	(0.060)
SEAT	1.947	(0.100)	Nissan	-0.269	(0.058)
Skoda	2.730	(0.098)	Opel	-0.216	(0.052)
smart	3.22	(0.149)	Peugeot	-0.170	(0.054)
Subaru	0.166	(0.091)	Renault	-0.250	(0.053)
Suzuki	0.96	(0.098)	SEAT	-0.249	(0.052)
Tesla	1.789	(0.465)	Skoda	-0.228	(0.053)
Toyota	1.175	(0.090)	smart	-0.055	(0.102)
Volvo	1.419	(0.090)	Subaru	-0.036	(0.056)
VW	2.971	(0.087)	Suzuki	-0.200	(0.057)
Compact Executive	0.353	(0.099)	Tesla	-0.243	(0.088)
Executive	0.728	(0.155)	Toyota	-0.023	(0.053)
Luxury	1.843	(0.225)	Volvo	-0.088	(0.053)
Mid-size	0.123	(0.049)	VW	-0.133	(0.052)
Coupe	-1.403	(0.139)	Compact Executive	0.270	(0.028)
Station wagon	1.351	(0.137)	Executive	0.259	(0.040)
Roadster	-1.109	(0.140)	Luxury	0.459	(0.048)
Hatchback	1.358	(0.132)	Mid-size	0.159	(0.017)
Sedan	-0.717	(0.135)	Coupe	-0.187	(0.029)

Table B.1: Demand and marginal cost estimates *(continued)*

	Coefficient	SE		Coefficient	SE
SUV	1.742	(0.120)	Station wagon	-0.246	(0.023)
Van	1.365	(0.128)	Roadster	0.054	(0.041)
ber	-1.079	(0.076)	Hatchback	-0.303	(0.024)
bra	-0.741	(0.070)	Sedan	-0.260	(0.028)
bre	-1.505	(0.123)	SUV	-0.146	(0.024)
bwt	-0.923	(0.094)	Van	-0.214	(0.025)
ham	-0.417	(0.073)			
hes	-0.320	(0.077)			
mvp	-0.608	(0.064)			
nie	-1.064	(0.078)			
nrw	-0.916	(0.094)			
rlp	-0.857	(0.090)			
sac	-0.372	(0.056)			
san	-1.012	(0.083)			
sar	-0.158	(0.058)			
swh	-1.002	(0.092)			
thr	-0.451	(0.058)			
Interactions					
Price / Income	-5.713	(0.691)			
Standard Dev.					
BHEV	2.563	(0.685)			
Fuel Cost	0.134	(0.017)			

Note:

Prices deflated and in EUR 1,000. Vehicle class-, Body-, Firm- and State Fixed Effects included.

Appendix C

Robustness to alternative corrections

Table C.1 shows estimates of key demand parameters under different corrections for observations with zero market shares. The column *Min bias* holds the results from the correction employed in the paper that follows D'Haultfœuille et al. (2019). The second column (*Laplace*) uses a correction based on Laplace's rule of succession that is used in Gandhi, Lu, and Shi (2013). It consists of replacing market shares by $s_{\tilde{j}mt} = \frac{\mathcal{M}_{mt}s_{jmt}+1}{\mathcal{M}_{mt}s_{jmt}+J_{mt}+1}$, with J_{mt} the number of products in market mt. Finally, column 3 (*Naive*) uses a naive correction where quantities of zero sales observations are assumed to be 1. We can see that the estimates do not change dramatically across the different corrections.

Table C.1: Estimates of key parameters under alternative corrections for zero market shares

	Min bias	**Laplace**	**Naive**
Mean Utility			
Range	1.772	1.694	1.792
	(0.223)	(0.193)	(0.213)
Range x Trend	-0.118	-0.105	-0.116
	(0.024)	(0.023)	(0.023)
Charging Stations	0.61	0.441	0.584
	(0.156)	(0.151)	(0.153)
Fuel Cost	-0.281	-0.263	-0.279
	(0.027)	(0.024)	(0.026)
BEV	-8.285	-5.955	-7.976
	(1.539)	(2.372)	(1.593)
PHEV	-5.901	-3.704	-5.598
	(1.482)	(2.353)	(1.544)
Interactions			
Price / Income	-5.713	-4.172	-5.275
	(0.691)	(0.548)	(0.659)
Standard Dev.			
EV	-2.563	-1.294	-2.368
	(0.685)	(1.798)	(0.743)
Fuel Cost	0.134	0.122	0.13
	(0.017)	(0.015)	(0.017)

Note: Standard errors in parentheses.

Appendix D

Estimated price elasticities in selected papers

Table D.1 presents estimates of price elasticities from several papers using a similar structural model of demand to mine.

Table D.1: Estimated price elasticities of selected papers

Author(s)	Price elasticity
Beresteanu and Li (2011)	-10.91
Berry et al. (1995)[1]	-3.928
Berry et al. (1995)[2]	-3.461
Li (2019)	-2.732
Klier and Linn (2012)	-2.6
Pavan (2017)	-2.85
Reynaert and Sallee (Forthcoming)	-5.45
Springel (2020)[3]	[-1, -1.5]
Thurk (2018)	-3.6

Own estimated price elasticity: -3.267

[1] Conlon and Gortmaker (Forthcoming) replication

[2] Conlon and Gortmaker (Forthcoming) own procedure

[3] Range of elasticities for EVs

Appendix E

A model of quality provision

E.1 Monopoly

In this section, I outline a model of quality provision by a monopolist. This model helps to understand the forces that determine how price and quality adjust to the introduction of a subsidy or a decrease in the marginal cost of quality provision. Note that what I call quality in this model can, in principle, be any product characteristics, such as driving range.

Set-up

Let us consider a monopolist who chooses price (p) and quality (q) of a single product sold to final consumers.[1] In my application, q would be the driving range of a car. The demand function $s(p, q)$ is increasing in quality, decreasing in price, and twice differentiable. Cost is an increasing function of quality and is denoted $c(q)s(p, q)$. A social planner subsidizes the product with a subsidy denoted by λ, possibly to increase the diffusion of the product. This scheme mirrors the type of subsidy for electric vehicles employed in countries such as Germany.

Quality choice

The monopolist maximizes its total profits given by $\pi(p, q)$. His optimization problem is given by

$$\max_{p,q} \pi(p, q) \equiv \ (p + \lambda - c(q)) \ s(p, q)$$

[1] The set-up slightly differs from Spence (1975) and Sheshinski (1976) where the monopolist's choice variables are quality and quantity.

and the first-order conditions of the monopolist are given by

$$\begin{aligned}
\text{[p]:} \quad & \pi_p \equiv s(p,q) + (p + \lambda - c)\frac{\partial s(p,q)}{\partial p} = 0 \\
\text{[q]:} \quad & \pi_q \equiv -c_q s(p,q) + (p + \lambda - c)\frac{\partial s(p,q)}{\partial q} = 0.
\end{aligned}$$

For the price, we recover the standard optimal markup formula. For quality, the formula looks similar. The firm faces a trade-off: It can increase quality to expand sales. However, doing so is costly and leads to a smaller margin. To see how the monopolist chooses quality in equilibrium, we can plug the price FOC into the quality FOC and re-arrange to find

$$c_q = \frac{\partial s(p,q)/\partial q}{|\partial s(p,q)/\partial p|}, \tag{E.1}$$

where c_q is the marginal cost of providing quality $\frac{\partial c(q)}{\partial q}$ The monopolist sets quality such that the marginal cost of providing quality is equal to the absolute value of the ratio of semi-elasticities of quality and price. The larger the fraction on the right-hand side of equation (E.1), the larger the level of quality provided in equilibrium.

The effect of a subsidy

What happens when the policymaker introduces a subsidy? If quality cannot adjust, we expect the monopolist to pass on the subsidy by lowering the price. The extend of this pass-through depends on the curvature of the demand curve. The more elastic the demand curve, the higher the amount of pass-through. If both the price and quality can adjust, there is no clear-cut answer to how the monopolist will react. Differentiating the system of first order conditions gives

$$\begin{bmatrix} \frac{dp}{d\lambda} \\ \frac{dq}{d\lambda} \end{bmatrix} = \begin{bmatrix} \pi_{pp} & \pi_{pq} \\ \pi_{pq} & \pi_{qq} \end{bmatrix}^{-1} \begin{bmatrix} -\pi_{p\lambda} \\ -\pi_{q\lambda} \end{bmatrix},$$

where π_{mn} denotes the second order derivative of the monopolist's profit function respect to m and n, with $m, n \in \{p, q\}$ and where

$$\begin{aligned}
\pi_{pp} &= 2s_p + s_{pp}(p + \lambda - c) \\
\pi_{qq} &= -c_{qq}s - 2c_q s_q + s_{qq}(p + \lambda - c) \\
\pi_{pq} &= s_q + (p + \lambda - c)s_{pq} - c_q s_p \\
\pi_{p\lambda} &= s_p, \quad \pi_{q\lambda} = s_q.
\end{aligned}$$

This gives

$$\frac{dp}{d\lambda} = \frac{1}{\Delta}\Big(\pi_{pq}\pi_{q\lambda} - \pi_{qq}\pi_{p\lambda}\Big)$$
$$\frac{dq}{d\lambda} = \frac{1}{\Delta}\Big(\pi_{pq}\pi_{p\lambda} - \pi_{pp}\pi_{q\lambda}\Big),$$

where $\Delta \equiv \pi_{pp}\pi_{qq} - \pi_{pq}^2 > 0$ from the second order conditions of having a global maximum. The SOCs further require $\pi_{pp} < 0$ and $\pi_{qq} < 0$. Note that we also have $\pi_{p\lambda} < 0$ and $\pi_{q\lambda} > 0$. If $\pi_{pq} < 0$, meaning price and quality are strategic substitutes, we have $\frac{dp}{d\lambda} < 0$ and $\frac{dq}{d\lambda} > 0$. In the case where $\pi_{pq} > 0$, things become more ambiguous. Note that we can write

$$\frac{dp}{d\lambda} = \frac{1}{\Delta}\Big(\pi_{pq}s_q - \pi_{qq}s_p\Big)$$
$$\frac{dq}{d\lambda} = \frac{1}{\Delta}\Big(\pi_{pq}s_p - \pi_{pp}s_q\Big),$$

We can then conclude that

$$\text{sign}\Big(\frac{dp}{d\lambda}\Big) = \text{sign}\Big(\Big|\frac{s_q}{\pi_{qq}}\Big| - \Big|\frac{s_p}{\pi_{pq}}\Big|\Big)$$
$$\text{sign}\Big(\frac{dq}{d\lambda}\Big) = \text{sign}\Big(\Big|\frac{s_p}{\pi_{pp}}\Big| - \Big|\frac{s_q}{\pi_{pq}}\Big|\Big)$$

The effect of a subsidy on quality and price depends on the relative magnitudes of the price- and quality semi-elasticities, s_p and s_q, and the marginal cost of providing quality c_q. Moreover, we can rule out the case $\pi_{p\lambda} > 0$ and $\pi_{q\lambda} < 0$. To see see why, note that this case would imply $\frac{\pi_{pq}}{\pi_{pp}} < \frac{s_q}{s_p} < \frac{\pi_{qq}}{\pi_{pq}}$ which violates the second order conditions.

E.2 Multi-product oligopoly

In this section I show how the main insights obtained in the monopoly case generalize to a multi-product oligopoly setting. The fact that there are cannibalization effects within a firm's product portfolio and the fact that products are differentiated within and across the product portfolio will influence the effect of a subsidy on price and quality but not alter the main conclusions. To see why, let us consider the following setting: There are $j = 1, \dots J$ products in a market. Consumers care about the quality of a subset of products $j \in \mathcal{B}$ and do not have any preferences over the quality of the remaining products $j \in \mathcal{I}$.[2] The social planner puts a subsidy on products in $\mathcal{B}$ but not on those in $\mathcal{I}$. Let us look at the firm f's profit maximization

[2]Think of the market for cars: The range of electric cars is a proxy for quality and costly to provide. Consumers do not care about the range of diesel or gasoline cars as it is sufficiently high and firms do not give it first-order importance when making their strategic decisions.

problem:

$$\max_{p_f,q_f} \pi_f = \sum_{k\in\mathcal{J}_f\cap k\in\mathcal{B}} (p_k+\lambda-c(q_k))s_k(p,q) + \sum_{l\in\mathcal{J}_f\cap k\in\mathcal{I}} (p_l-c(q_l))s_l(p,q),$$

where p_f and q_f denote the own-firm vectors of price and quality, respectively, p and q the price-and quality vectors of all firms in the market and J_f the portfolio of firm-f products. The FOCs for product one are then given by

$$\begin{aligned}
[p_1]: \quad & \pi_{fp_1} \equiv s_1 + \sum_{k\in\mathcal{J}_f\cap k\in\mathcal{B}} (p_k+\lambda-c(q_k))\frac{\partial s_k}{\partial p_1} + \sum_{l\in\mathcal{J}_f\cap k\in\mathcal{I}} (p_l-c(q_l))\frac{\partial s_l}{\partial p_1} = 0\\
[q_1]: \quad & \pi_{fq_1} \equiv -c_{q_1}s_1 + \sum_{k\in\mathcal{J}_f\cap k\in\mathcal{B}} (p_k+\lambda-c(q_k))\frac{\partial s_k}{\partial q_1} + \sum_{l\in\mathcal{J}_f\cap k\in\mathcal{I}} (p_l-c(q_l))\frac{\partial s_l}{\partial q_1} = 0
\end{aligned}$$

The second-order derivatives of the profit function will depend not only on the effect of own price and quality on own demand, but also on the demand of the other own-firm products. Finally, they depend on rival product prices and quantities through the demand function.

Increase of subsidy for a single product

In the case where the subsidy is only increased for a single product product, say product 1, we get

$$\begin{aligned}
\frac{dp_1}{d\lambda} &= \frac{1}{\Delta}\Big(\pi_{fp_1q_1}\pi_{fq_1\lambda} - \pi_{fq_1q_1}\pi_{fp_1\lambda}\Big)\\
\frac{dq_1}{d\lambda} &= \frac{1}{\Delta}\Big(\pi_{fp_1q_1}\pi_{fp_1\lambda} - \pi_{fp_1p_1}\pi_{fq_1\lambda}\Big),
\end{aligned}$$

meaning that the general results from above go through: The signs of $\frac{dp_1}{d\lambda}, \frac{dq_1}{d\lambda}$ depend on whether p, q are strategic substitutes or complements. They also still depend on the marginal cost of providing quality as well as the relative magnitudes of $\pi_{fp_1\lambda}$ and $\pi_{fq_1\lambda}$ that themselves still depend on s_p and s_q.

Increase in the subsidy for all products in $\mathcal{B}$

Things become more complicated when we consider an increase on the subsidy of all products in $\mathcal{B}$. We now need to differentiate $J + J_{\mathcal{B}}$ first order conditions ($J_{\mathcal{B}}$ being the cardinality of $\mathcal{B}$). In essence, the effect of price and quality on the FOC of all other products now needs to be taken into account as well.

Let J denote the cardinality of all products, $J_{\mathcal{B}}$ the cardinality of those products with endogenous quality and $f(j)$ the firm of product j. Then, we have the following system of FOCs with

$J + J_q$ equations:

$$[p_1]: \quad \pi_{f(1)p_1} \equiv s_1 + \sum_{k \in \mathcal{J}_{f(1)} \cap k \in \mathcal{B}} (p_k + \lambda - c_k) \frac{\partial s_k}{\partial p_1} + \sum_{l \in \mathcal{J}_{f(1)} \cap l \in \mathcal{I}} (p_l - c_l) \frac{\partial s_l}{\partial p_1} = 0$$

$$\vdots$$

$$[p_J]: \quad \pi_{f(J)p_J} \equiv s_J + \sum_{k \in \mathcal{J}_{f(1)} \cap k \in \mathcal{B}} (p_k + \lambda - c_k) \frac{\partial s_k}{\partial p_J} + \sum_{l \in \mathcal{J}_{f(1)} \cap l \in \mathcal{I}} (p_l - c_l) \frac{\partial s_l}{\partial p_J} = 0$$

$$[q_1]: \quad \pi_{f(1)q_1} \equiv -c_{q_1} s_1 + \sum_{k \in \mathcal{J}_{f(1)} \cap k \in \mathcal{B}} (p_k + \lambda - c_k) \frac{\partial s_k}{\partial q_1} + \sum_{l \in \mathcal{J}_{f(1)} \cap l \in \mathcal{I}} (p_l - c_l) \frac{\partial s_l}{\partial q_1} = 0$$

$$\vdots$$

$$[q_{J_\mathcal{B}}]: \quad \pi_{f(J_\mathcal{B})q_{J_\mathcal{B}}} \equiv -c_{q_{J_\mathcal{B}}} s_{J_\mathcal{B}} + \sum_{k \in \mathcal{J}_{f(J_\mathcal{B})} \cap k \in \mathcal{B}} (p_k + \lambda - c_k) \frac{\partial s_k}{\partial q_{J_\mathcal{B}}} + \sum_{l \in \mathcal{J}_{f(J)} \cap l \in \mathcal{I}} (p_l - c_l) \frac{\partial s_l}{\partial q_{J_\mathcal{B}}} = 0$$

The total differentiation of this system yields

$$\begin{bmatrix} \frac{dp_1}{d\lambda} \\ \vdots \\ \frac{dp_J}{d\lambda} \\ \frac{dq_1}{d\lambda} \\ \vdots \\ \frac{dq_{J_\mathcal{B}}}{d\lambda} \end{bmatrix} = \begin{bmatrix} \pi_{f(1)p_1p_1} & \cdots & \pi_{f(J)p_Jp_1} & \pi_{f(1)q_1p_1} & \cdots & \pi_{f(J_\mathcal{B})q_{J_\mathcal{B}}p_1} \\ \vdots & \vdots & \vdots & \vdots & & \\ \pi_{f(1)p_1p_J} & \cdots & \pi_{f(J)p_Jp_J} & \pi_{f(1)q_1p_J} & \cdots & \pi_{f(J_\mathcal{B})q_{J_\mathcal{B}}p_J} \\ \pi_{f(1)p_1q_1} & \cdots & \pi_{f(J)p_Jq_1} & \pi_{f(1)q_1q_1} & \cdots & \pi_{f(J_\mathcal{B})q_{J_\mathcal{B}}q_1} \\ \vdots & \vdots & \vdots & \vdots & & \\ \pi_{f(1)p_1q_{J_\mathcal{B}}} & \cdots & \pi_{f(J)p_Jq_{J_\mathcal{B}}} & \pi_{f(1)q_1q_{J_\mathcal{B}}} & \cdots & \pi_{f(J_\mathcal{B})q_{J_\mathcal{B}}q_{J_\mathcal{B}}} \end{bmatrix}^{-1} \begin{bmatrix} -\pi_{f(1)p_1\lambda} \\ \vdots \\ -\pi_{f(J)p_J\lambda} \\ -\pi_{f(1)q_1\lambda} \\ \vdots \\ -\pi_{f(J_\mathcal{B})q_{J_\mathcal{B}}\lambda} \end{bmatrix}, \qquad \text{(E.2)}$$

where for instance

- $\pi_{f(1)p_1p_1} = 2\frac{\partial s_1}{\partial p_1} + \sum_{k \in \mathcal{J}_{f(1)} \cap k \in \mathcal{B}} (p_k + \lambda - c_k) \frac{\partial^2 s_k}{\partial p_1^2} + \sum_{l \in \mathcal{J}_{f(1)} \cap l \in \mathcal{I}} (p_l - c_l) \frac{\partial^2 s_l}{\partial p_1^2}$
- $\pi_{f(J)p_Jp_1} = \frac{\partial s_J}{\partial p_1} + \frac{\partial s_J}{\partial p_1} \mathbf{1}\{1, J \in f(J)\} + \sum_{k \in \mathcal{J}_{f(J)} \cap k \in \mathcal{B}} (p_k + \lambda - c_k) \frac{\partial^2 s_k}{\partial p_J \partial p_1} + \sum_{l \in \mathcal{J}_{f(J)} \cap l \in \mathcal{I}} (p_l - c_l) \frac{\partial^2 s_l}{\partial p_J \partial p_1}$
- $\pi_{f(1)p_1q_1} = -c_{q_1} \frac{\partial s_1}{\partial p_1} + \frac{\partial s_1}{\partial q_1} + \sum_{k \in \mathcal{J}_{f(1)} \cap k \in \mathcal{B}} (p_k + \lambda - c_k) \frac{\partial^2 s_k}{\partial p_1 \partial q_1} + \sum_{l \in \mathcal{J}_{f(1)} \cap l \in \mathcal{I}} (p_l - c_l) \frac{\partial^2 s_l}{\partial p_1 \partial q_1}$
- $\pi_{f(1)p_1q_{J_\mathcal{B}}} = -c_{q_{J_\mathcal{B}}} \frac{\partial s_{J_\mathcal{B}}}{\partial p_1} \mathbf{1}\{1, J_\mathcal{B} \in f(1)\} + \frac{\partial s_1}{\partial q_{J_\mathcal{B}}} + \sum_{k \in \mathcal{J}_{f(1)} \cap k \in \mathcal{B}} (p_k + \lambda - c_k) \frac{\partial^2 s_k}{\partial p_1 \partial q_{J_\mathcal{B}}} + \sum_{l \in \mathcal{J}_{f(1)} \cap l \in \mathcal{I}} (p_l - c_l) \frac{\partial^2 s_l}{\partial p_1 \partial q_{J_\mathcal{B}}}$
- $\pi_{f(1)q_1q_1} = -c_{q_1q_1} s_1 - 2c_{q_1} \frac{\partial s_1}{\partial q_1} + \sum_{k \in \mathcal{J}_{f(1)} \cap k \in \mathcal{B}} (p_k + \lambda - c_k) \frac{\partial^2 s_k}{\partial q_1^2} + \sum_{l \in \mathcal{J}_{f(1)} \cap l \in \mathcal{I}} (p_l - c_l) \frac{\partial^2 s_l}{\partial q_1^2}$
- $\pi_{f(1)q_1q_{J_\mathcal{B}}} = -c_{q_{J_\mathcal{B}}} \frac{\partial s_{J_\mathcal{B}}}{\partial q_1} \mathbf{1}\{1, J_\mathcal{B} \in J_f\} - c_{q_1} \frac{\partial s_1}{\partial q_{J_\mathcal{B}}} + \sum_{k \in \mathcal{J}_{f(1)} \cap k \in \mathcal{B}} (p_k + \lambda - c_k) \frac{\partial^2 s_k}{\partial q_1 \partial q_{J_\mathcal{B}}} + \sum_{l \in \mathcal{J}_{f(1)} \cap l \in \mathcal{I}} (p_l - c_l) \frac{\partial^2 s_l}{\partial q_1 \partial q_{J_\mathcal{B}}}$
- $\pi_{p_1\lambda} = \sum_{k \in \mathcal{J}_{f(1)} \cap k \in \mathcal{B}} \frac{\partial s_k}{\partial p_1}$

- $\pi_{q_1\lambda} = \sum_{k\in\mathcal{J}_{f(1)}\cap k\in\mathcal{B}} \frac{\partial s_k}{\partial q_1}$

It is no longer possible to simply pin down the effects of the subsidy on whether or not p, q are strategic complements, nor on the relative magnitudes of $\pi_{fp_1\lambda}$ and $\pi_{fq_1\lambda}$ and the marginal cost of providing quality. First off however, the entries $\pi_{fp_jp_j}$ and $\pi_{fq_jq_j}$ in the matrix to be inverted in E.2 are likely to dominate the entries $\pi_{fp_jp_k}$ and $\pi_{fq_jq_k}$, $k \neq j$. Hence the signs and magnitudes of these own second-order derivatives will play an important role in determining the effect of the subsidy. Secondly, the system in E.2, while too opaque to be solved analytically, can be solved numerically if estimated profits and semi-elasticities can be recovered and prices as well as qualities are known. I can do so in my empirical setting below. In principle, this system can also be obtained to measure pass-through of a change in marginal cost. The difference is then that the system of first order conditions will be differentiated with respect to the change in marginal cost. Finally, the case where several multi-product firms produce products with endogenous quality that are subsidized and products with fixed quality that are not subsidized. Note that a similar system can be obtained to analyze pass-through of a shock to the marginal cost of providing quality.

Chapter 2

Price Competition and Endogenous Product Choice in Networks: Evidence from the US Airline Industry

WITH CHRISTIAN BONTEMPS AND CRISTINA GUALDANI

2.1 Introduction

The airline industry has received a lot of attention in the economic literature, sparked by the wave of mergers and bankruptcies after the U.S. Airline Deregulation Act of 1978.[1] Most of the contributions on merger and bankruptcy evaluations are based on supply-demand models, where the airlines best respond to competitors by adjusting their prices, while holding the entry decisions in markets exogenous and fixed (Bresnahan, 1987; Berry, 1994; Berry, Levinsohn, and Pakes, 1995). More plausibly, such events prompt the airlines to best respond also by repositioning in markets. For example, a merger could generate cost savings for the merged firm which may favour its entry in new markets. Also, after a merger, there might room in some markets for accommodating other entrants. The aim of this paper is to provide a tractable framework for the airline industry which combines entry and pricing decisions and use it to conduct counterfactual exercises where the airlines are allowed to modify prices and market structures.

More broadly, the question about endogenising entry decisions (and, hence, product choices and characteristics) in supply-demand models is of general interest and it was already posed in earlier works. For example, Berry (1994): "*I should emphasize in closing that the techniques of this article rely on a number of restrictive assumptions.* [...] *More importantly, and more difficult to solve, I assume that product characteristics are econometrically exogenous.*" (p.260). Berry et al. (1995): "*The second, and richer, part of the problem is to endogenize the actual choice of the characteristics of the models marketed.*" (p.886). Berry and Jia (2010): "*An implicit assumption of our empirical model is that the network structure and the carriers that serve each market are taken as given. Ideally, we would like to model a three-stage game. First, carriers form their hubs. Second, given the hub structure, each carrier chooses a set of markets to serve. Third, given these entry decisions, carriers compete in prices and the frequency of flight departures.*" (p.20).

Endogenising entry decisions in a supply-demand model for the airline industry is challenging. This is because the entry decisions of an airline are interdependent across markets. Specifically, the literature on the airline industry suggests that hub-and-spoke operations decrease the marginal costs of serving markets out of hubs, via economies of density and scope, and increase the total fixed costs, due to congestion at hubs. Further, customers may find it attractive to fly from dense hubs because this enhances the value of frequent flier programs.[2] Such synergies among markets imply that the entry decision of an airline in market t may *spill over*

[1] See Table H.0.1 in Appendix H.

[2] See, for example, Caves, Christensen, and Tretheway (1984), Kanafani and Ghobrial (1985), Morrison and Winston (1986), Levine (1987), Butler and Houston (1989), Berry (1990),Borenstein (1989; 1992), Butler and Houston (1989), Morrison and Winston (1989),Berry (1990), Brueckner, Dyer, and Spiller (1992), Brueckner and Spiller (1994), Oum, Zhang, and Zhang (1995), Berry, Carnall, and Spiller (1996), Nero (1999), Berry and Jia (2010), and Berry, Gaynor, and Scott Morton (2019).

its entry decision in market t' by directly affecting the demand, marginal cost, and fixed cost equations (and, in turn, profitability) in market t'. Due to these *spillover effects*, an airline does not take its entry decisions independently across markets. Instead, an airline globally forms the *network* of markets to be served in order to internalise the spillover effects within its optimisation problem.[3]

This paper addresses the above concerns by developing an empirical two-stage game. In the first stage, the airlines form their networks by entering markets and pay the fixed costs. The network an airline consists of a collection of nodes corresponding to the cities (or, airports) and a collection of links between nodes representing the markets served by the airline.[4] In the second stage, the airlines face the demand for their products, pay the variable costs, and choose which prices to charge by competing in a standard Bertrand-Nash pricing game. Thus, in the first stage, the airlines form the networks which maximise the expected second-stage profits net of the fixed costs, while taking into account spillover effects in entry decisions across markets on the demand, marginal cost, and fixed costs sides, as discussed in the previous paragraph.

Identifying the parameters governing our two-stage game is challenging. While identification of the second-stage parameters can be established by following the standard approach for supply-demand models with differentiated products (Berry and Haile, 2014), identification of the first-stage parameters is complicated by the discrete choice nature of the problem. In particular, there are two main issues. First, there may be multiple Nash equilibrium networks. This is because the airlines compete at the entry stage through the second-stage pricing game. In turn, we are not able to write down a well-defined likelihood function. Second, even if one is willing to specify an equilibrium selection mechanism, it remains burdensome to construct the set of Nash equilibrium networks, due to the large number of markets. In turn, we are not able to write down a tractable likelihood function that can be evaluated many times throughout an optimisation routine. Instead of focusing on Nash equilibrium networks, we bypass these two issues by considering *implications* of (i.e., *necessary* conditions for) Nash equilibrium, which are easier to handle from an econometric point of view. More precisely, we implement the revealed preference approach by Pakes (2010) and Pakes, Porter, Ho, and Ishii (2015). This approach consists of two steps. First, we write down inequalities by simply predicting that the observed networks lead to higher profits than the profits would be were the airlines to deviate from the observed networks. Second, we get rid of the structural errors entering the inequalities by taking the expectation of the inequalities over markets and interacting them with appropriate instruments. The resulting moment inequalities characterise a non-sharp identified set that is

[3]The importance of network considerations has been also highlighted in several merger investigations. For example, see the Department of Justice's Competitive Impact statement on the merger between American Airlines and US Airways (`https://www.justice.gov/atr/case-document/file/514516/download`).

[4]For example, Figure H.0.1 in Appendix H represents the network of markets served by American Airlines, before the merger with US Airways.

a convex polytope, whose projections can be easily obtained by solving linear programming problems.

We estimate the model using data from the US Airline Origin and Destination Service, a 10% random sample of all tickets issued in the United States during the second quarter of 2011. We focus on flights operated between the 85 largest metropolitan statistical areas in the United States, which are served by United Airlines, Delta Airlines, American Airlines, US Airways, Southwest Airlines, low and medium cost carriers. In the first stage, we find that fixed costs increase in the number of destinations reachable from hub airports. On the supply side of the second stage, we find that marginal costs decrease in the number of flights (direct or one-stop) offered out of the endpoints. On the demand side of the second stage, we find that consumer utility increases in the number of direct connections that can be reached from the endpoints. As mentioned above, previous works have suggested that the size of hub-and-spoke networks increase the fixed costs and decrease the marginal cost. To the best of our knowledge, our model is the first to confirm these effects in a structural model of the airline market.

With estimates of the two stages in hand, we simulate the effects of a merger between American Airlines and US Airways and compare the merger, which did occur in 2013, to a bankruptcy and subsequent disappearance of American Airlines.[5] In a first step, we compare the predictions based on our full model to the predictions form a model in which networks do not adjust in response to a merger or bankruptcy. We find that leaving the network unchanged or making ad-hoc assumptions about it can lead to misleading conclusions regarding market outcomes. In a second step, we investigate the merger and, in particular, the effect of remedies that forced the merged entity to not reduce operations at most of its hubs. We find that these remedies turned a slight decrease in consumer surplus into a slight increase. At the most negatively affected hubs, the remedies helped to contain harm to consumers. In a third step, we compare the merger to a scenario in which a bankruptcy of American Airlines leads to its disappearance from the market. Not surprisingly, the bankruptcy induces more rival firm entry than the merger. However, the loss of access to a large network causes substantial loss in consumer surplus at American's hubs. Other firms are not able to fill this void completely. Overall, consumer surplus decreases by more than in the merger case. Our results underline one key advantage of a conditional merger over the bankruptcy of a distressed hub-and-spoke airline: competition authorities can shape post-merger outcomes by imposing remedies.

[5] When the two airlines announced their intention to merge, American Airlines was under Chapter 11 bankruptcy.

2.2 Literature review

This paper aims to bridge a gap between two strands of the literature. The first strand is the literature on supply-demand models. This literature estimates demand and supply equations, while taking entry decisions as exogenously given (for example, Bresnahan, 1987; Berry, 1994; Berry et al., 1995; Berry, Linton, and Pakes, 2004; Berry and Haile, 2014). Some empirical studies using supply-demand models for the airline industry are Berry et al. (1996), Berry and Jia (2010), Ciliberto and Williams (2014), Peters (2006), Mark, Keating, Rubinfeld, and Willig (2013), and Das (2019). The second strand is the literature on entry models (for example, Reiss and Spiller, 1989; Bresnahan and Reiss, 1990; 1991; Berry, 1992; Goolsbee and Syverson, 2008; Ciliberto and Tamer, 2009). This literature estimates the payoffs from entering markets, under the assumption that entry decisions are independent across markets and without considering demand and supply. In this paper, we model entry decisions, supply, and demand. Moreover, we allow for spillover effects in entry decisions across markets, as discussed in Section 2.1.

There are other papers which combine entry and pricing decisions for studying the airline industry. The seminal contributions by Li, Mazur, Park, Roberts, Sweeting, and Zhang (2019), and Ciliberto, Murry, and Tamer (2021) acknowledge the network dimension, but treat spillover effects as exogenous product covariates and assume that entry decisions are independent across markets. In turn, when simulating a merger, those approaches require one to focus *separately* on each market and ad-hoc readjust the spillover effects, for example by distinguishing a base-case from a best-case scenarios. This is not needed in our framework because the airlines are allowed to reoptimise their entire networks. In turn, we can offer a *global* view of the overall changes in the networks and other market outcomes. Aguirregabiria and Ho (2012) develop a dynamic game of entry and pricing decisions, but assume that entry decisions are decentralised at the market level (i.e., each market is run by a local manager taking independent entry decisions), which substantially reduces the dimensionality of the strategy space. Here, instead, we model network formation as a process centralised at the level of each airline. Also Benkard, Bodoh-Creed, and Lazarev (2020) consider a dynamic setting, but do not include the demand equation. Hence, differently from our paper, their framework cannot be used to quantify welfare effects. Admittedly, though, our model does not incorporate dynamics, which we leave to future extensions. Park (2020) combine entry and pricing decisions, but entry is endogenised only in the markets out of Ronald Reagan Washington National Airport. Here, instead, we endogenise entry in all markets. Lastly, Yuan (2020) is close in spirit to our approach by developping a three-stage game of entry, frequency, and price choices. However, he obtains point estimates for the first-stage parameters via a calibration strategy and does not consider spillover effects in entry decisions across markets on the fixed cost side.

A number of papers combine entry and pricing decisions for studying other industries (for example, Eizenberg, 2014; Holmes, 2011; Houde, Newberry, and Seim, 2017; Kuehn, 2018; Rossetti, 2018; Wollmann, 2018). As most of these papers, we rely on the revealed preference approach by Pakes (2010) and Pakes et al. (2015) in order to simplify identification of the first-stage parameters. Here, however, we face spillover effects in entry decisions across markets which appear on the demand, marginal cost, and fixed cost sides and create lots of computational challenges.

Finally, our paper broadly connects with the recent advances in the econometrics of network formation (e.g., Chandrasekhar, 2016; Graham, 2015; de Paula, 2017; 2020; Graham and de Paula, 2020). We have decided to pursue the revealed preference approach by Pakes (2010) and Pakes et al. (2015), rather than applying those methods, for two reasons. First, the latter involve computationally serious challenges in the presence of spillover effects, which become even deeper when combined with our second stage. Instead, the moment inequalities that we obtain from the revealed preference approach are computationally easier to handle because linear in the first-stage parameters. Second, the latter typically view network formation as a process decentralised at the level of each node. Here, instead, network formation is centralised at the level of each airline.

The rest of the paper is organised as follows. Section 2.3 presents the model. Section 2.4 discusses identification. Section 2.5 describes the empirical application. Section 2.6 concludes. Further details are in the appendices.

2.3 The model

There are N airlines, labelled by $f \in \mathcal{N} \equiv \{1, \dots, \mathrm{N}\}$, which play a two-stage game. In the first stage, the airlines form their networks by entering markets and pay the fixed costs. In the second stage, the airlines face the demand for their products, pay the variable costs, and choose which prices to charge. In the first stage, the airlines observe their own and the competitors' fixed cost shocks. However, they do not observe their own and the competitors' demand and marginal cost shocks, despite knowing the probability distribution from which such shocks are drawn. Hence, in the first stage, the airlines form expectations about the second-stage profits. The airlines publicly discover the demand and marginal cost shocks just before playing the second stage, where they compete in a standard Bertrand-Nash pricing game. The airlines solve the game by going backward from the second stage and use pure strategy subgame perfect Nash equilibrium as solution concept. In what follows, we describe the game in more details starting from the second stage.[6]

[6]We do not model frequency and capacity choices. This is because we want to preserve tractability, given the many challenges introduced by the network formation stage. Moreover, modelling frequency and capacity choices would require us to collect detailed data on the types of aircrafts used, the flight schedule, and the

Figure 2.1: Timing of the game

2.3.1 The second stage

In the second stage, the airlines take as given the entry decisions in markets and the consequent product choices. Markets are undirectional city-pairs (for example, Boston-Houston).[7] Products are airline-itinerary combinations (for example, Boston-Houston via Miami operated by American Airlines).[8] In every market, the airlines face the demand for their products, pay the variable costs, and simultaneously choose the prices maximising the variable profits, under complete information. We use a standard supply-demand framework for differentiated products.

Demand We consider the Nested Logit demand with two nests, one for the airline products, the other for the outside option of not travelling or travelling with other means (Berry, 1994).[9] Each market is indexed by $t \in \mathcal{T}$, where $\mathcal{T}$ is the set of markets. Each product offered in market t is indexed by $j \in \mathcal{J}_t$, where $\mathcal{J}_t$ is the set of products offered in market t. The utility that individual i receives from buying in market t is specified as:

$$\begin{aligned} &\text{product } j: && U_{i,j,t} = X_{j,t}^{\top}\beta - \alpha P_{j,t} + \xi_{j,t} + \nu_{i,t} + \lambda\epsilon_{i,j,t}, \\ &\text{outside option 0:} && U_{i,0,t} = \epsilon_{i,0,t}. \end{aligned} \tag{2.1}$$

In (2.1), $X_{j,t}$ is a vector of product characteristics and $P_{j,t}$ is the product price. Both $X_{j,t}$ and $P_{j,t}$ are observed by the researcher. $\xi_{j,t}$ represents the product characteristics that are unobserved by the researcher but are observed by the airlines. $\nu_{i,t}, \epsilon_{i,j,t}, \epsilon_{i,0,t}$ are the consumer tastes, unobserved by the researcher, i.i.d. across i, j, t, and independent of all other variables. The probability distribution of the consumer tastes is chosen to yield the familiar Nested Logit market share function, with $\lambda \in (0, 1]$.

We include in $X_{j,t}$ various product characteristics, such as the number of stops, the distance flown and its squared value, and the number of direct flights offered at the itinerary's origin

number of passengers on each flight. We leave extensions on these aspects to future analysis.

[7] We do not distinguish between airports in the same city. In fact, carriers in nearby airports might compete against each other because customers can choose which airport to fly from. We do not define markets as directional city-pairs because our data contain very few cases of airlines not serving both directions of a given city-pair.

[8] Note that the airlines can be multi-product because, in a given market, an airline may offer direct flights and/or connecting flights. Further, two itineraries offered by an airline with the same endpoints but different connecting cities are treated as two different products.

[9] Also Ciliberto et al. (2021) adopt the Nested Logit demand for studying the airline industry.

by the same carrier offering itinerary j (hereafter, "Connections"[10]). We expect travellers to prefer direct flights. We also expect that, as distance increases, air travel becomes more attractive relative to the outside option. However, as distance increases further, travel becomes less pleasant and demand starts to decrease. The variable Connections captures the value of frequent flier programs. In fact, the larger the number of destinations for which consumers can redeem frequent flier miles, the higher the value of such loyalty programs. Additionally, an airline that flies to many cities is likely to have more convenient parking and gate access and provide better services. For a thorough discussion on the impact of hubbing on costumer utility, see Levine (1987), Borenstein (1989; 1992), Butler and Houston (1989), Morrison and Winston (1989), Berry (1990), Oum et al. (1995), Berry et al. (1996), and Berry and Jia (2010). We include in $X_{j,t}$ carrier fixed effects to control for brand preferences. We also add city fixed effects to catch unobserved heterogeneity at the city level, such as leading economic sectors, climate, and infrastructures.

Note that, due to the variable Connections, the demand for product j in market t depends on the entry decisions of an airline in other markets $t' \neq t$. This gives rise to *spillover effects* in entry decisions across markets on the *demand* side.

We compute $P_{j,t}$ as the average of the fares in the data sharing the same airline-itinerary combination, weighted by the number of passengers.[11] Alternatively, one could introduce a finite number of "fare bins" for each airline-itinerary combination, as in Berry and Jia (2010). We have decided not to do so because, in our two-stage setting, we find more reasonable to consider an average price that the carriers expect or wish to achieve for each offered itinerary. Further, having fare bins would add many methodological and computational challenges. For instance, we would need to carefully examine how to define fare bins in a way that does not make the empirical results excessively sensitive to such a definition. We would also need to deal with the fact that some fare-airline-itinerary combinations may not be available at all time to customers and/or may have zero market shares in the data.[12] Given that our model is already complicated by the network formation stage, we prefer to leave these extensions to future research. In turn, we do not allow for heterogeneity in consumer taste for price. This is because, with average prices, there is not sufficient price variation left in the data.

[10] More precisely, given that markets are defined as undirectional city-pairs, the variable Connections is computed as the maximum between the numbers of direct flights offered at the endpoints of itinerary j by the same carrier offering itinerary j. For example, suppose that market t is Boston-Houston and product j is a direct flight between Boston and Houston operated by American Airlines. Suppose that American Airlines offers direct flights to 3 destinations out of Boston and to 5 destinations out of Houston. Then, the variable Connections is equal to $\max\{3, 5\} = 5$.

[11] See Section 3.2 for more details on the computation of prices.

[12] See, for instance, Abaluck and Adams (2020) and Barseghyan, Coughlin, Molinari, and Teitelbaum (2020) about the identification of discrete choice models with latent choice sets. See Gandhi, Lu, and Shi (2019) about the estimation of the demand for differentiated products with zeroes in market share data.

The demand shock, $\xi_{j,t}$, captures product characteristics that are not in our data and that can be arbitrarily correlated with prices, such as refundable versus non-refundable tickets and the quality of in-flight service. We do not specify any parametric distribution for $\xi_{j,t}$. Instead, to establish point identification of the second-stage parameters, in Section 2.4.1 we will assume that the demand shocks are uncorrelated with the observed demand shifters, as standard in the literature on supply-demand models. In the same section, we will discuss the pros and cons of this assumption.

From utility maximising behaviour, we obtain the predicted product shares in market t:

$$\begin{aligned} &\text{product } j: && s_{j,t}(X_t, P_t, \xi_t; \theta_d) = \frac{\exp(\delta_{j,t}/\lambda)}{D_t} \frac{D_t^\lambda}{1 + D_t^\lambda}, \\ &\text{outside option 0:} && s_{0,t}(X_t, P_t, \xi_t; \theta_d) = \frac{1}{1 + D_t^\lambda}, \end{aligned} \tag{2.2}$$

where $D_t \equiv \sum_{j=1}^{J_t} \exp(\delta_{j,t}/\lambda)$, $\delta_{j,t} \equiv X_{j,t}^\top \beta - \alpha P_{j,t} + \xi_{j,t}$, $\theta_d \equiv (\beta, \alpha, \lambda)$, $X_t \equiv (X_{j,t}\ \forall j \in \mathcal{J}_t)$, $P_t \equiv (P_{j,t}\ \forall j \in \mathcal{J}_t)$, and $\xi_t \equiv (\xi_{j,t}\ \forall j \in \mathcal{J}_t)$. In turn, the predicted demand in market t is:

$$\begin{aligned} &\text{product } j: && s_{j,t}(X_t, P_t, \xi_t; \theta_d) \times M_t, \\ &\text{outside option 0:} && s_{0,t}(X_t, P_t, \xi_t; \theta_d) \times M_t, \end{aligned} \tag{2.3}$$

where M_t is the market size, observed by the researcher. We further assume that the researcher observes the true product shares, as standard in the literature.

Supply In the second stage, the airlines pay the variables costs, such as the costs of fuel, oil, aircraft maintenance, landing fees, and passenger fees. We consider a constant and linear marginal cost specification.[13] In particular, the marginal cost of offering product j in market t is specified as:

$$\text{MC}_{j,t} = W_{j,t}^\top \psi + \omega_{j,t}. \tag{2.4}$$

In the above expression, $W_{j,t}$ is a vector of marginal cost shifters that are observed by the researcher. $\omega_{j,t}$ represents the marginal cost shifters that are unobserved by the researcher but are observed by the airlines.

We include in $W_{j,t}$ various product characteristics, such as the distance flown and the number of cities that are reachable from the endpoints and intermediate stops of itinerary j with the same carrier offering itinerary j (hereafter, "Presence"[14]). We expect the marginal costs to increase

[13] A constant and linear marginal cost specification is assumed in many other studies of the airline industry, e.g., Berry and Jia (2010), Ciliberto and Williams (2014), Das (2019), Li et al. (2019).

[14] More precisely, the variable Presence is computed as the average number of destinations that are reachable (with direct or connecting flights) from the endpoints and intermediate stops of itinerary j with the same carrier offering itinerary j. For example, suppose that market t is Boston-Houston and product j is a flight between

with distance due to the use of oil and fuel. The variable Presence captures the impact of economies of density on the marginal costs, i.e., the fact that more densely travel segments tend to have lower unit costs due to engineering reasons. In particular, the larger the number of final destinations consumers can reach, the more the opportunities for an airline to pool passengers from several itineraries into the same planes, and the more an airline can efficiently use large aircrafts which generally have lower unit costs. At the same time, many connections may also cause congestion and increase the fixed costs, as discussed in Section 2.3.2. For a comprehensive analysis on the marginal cost savings induced by hub-and-spoke operations, see Caves et al. (1984), Kanafani and Ghobrial (1985), Morrison and Winston (1986), Butler and Houston (1989), Berry (1990), Brueckner et al. (1992), Brueckner and Spiller (1994), Oum et al. (1995), Berry et al. (1996), Nero (1999), and Berry and Jia (2010). Analogously to the demand side, we also include in $W_{j,t}$ carrier fixed effects and city fixed effects.

Note that, due to the variable Presence, the marginal cost of product j in market t depends on the entry decisions of an airline in other markets $t' \neq t$. This gives rise to *spillover effects* in entry decisions across markets on the *marginal cost* side.

The airlines simultaneously set the prices in each market in order to maximise the variable profits, under complete information:

$$\text{variable profits of airline } f\text{:} \quad \sum_{t \in \mathcal{T}} \sum_{j \in J_{f,t}} (P_{j,t} - \text{MC}_{j,t}) \times s_{j,t}(X_t, P_t, \xi_t; \theta_d) \times M_t, \tag{2.5}$$

where $\mathcal{J}_{f,t}$ is the set of products offered by airline f in market t. For each airline f and market t, we obtain the Bertrand-Nash F.O.C.s in the usual way:

$$\text{MC}_{f,t} = P_{f,t} + \left(\frac{\partial s_{f,t}(X_t, P_t, \xi_t; \theta_d)}{\partial P_{f,t}} \right)^{-1} s_{f,t}(X_t, P_t, \xi_t; \theta_d), \tag{2.6}$$

where $\text{MC}_{f,t}$, $P_{f,t}$, and $s_{f,t}(X_t, P_t, \xi_t; \theta_d)$ are the vectors stacking $\text{MC}_{j,t}$, $P_{j,t}$, and $s_{j,t}(X_t, P_t, \xi_t; \theta_d)$, respectively, for each product $j \in \mathcal{J}_{f,t}$. $\frac{\partial s_{f,t}(X_t,P_t,\xi_t;\theta_d)}{\partial P_{f,t}}$ is the matrix collecting the partial derivatives of product shares with respect to prices. Note that, despite the airlines choose the prices market-by-market, the prices in a market indirectly depends on the prices in other markets due to the presence of spillover effects in entry decisions across markets.

2.3.2 The first stage

In the first stage, the airlines simultaneously form their networks by entering markets and pay the fixed costs. The airlines know everything about the second stage except their own and

Boston and Houston with an intermediate stop at Miami operated by American Airlines. Suppose that American Airlines allows to reach 3 destinations out of Boston, 5 destinations out of Miami, and 4 destinations out of Houston. Then, the variable Presence is equal to $(3 + 5 + 4)/3 = 4$.

the competitors' demand and marginal costs shocks. This is a natural assumption because the legacy carriers - which are the main players of our empirical application - typically operate with a time-lag between the entry decisions and the sale of flight tickets. Nevertheless, the airlines are aware of the probability distribution from which the second-stage shocks are drawn. Therefore, they can compute the expected second-stage profits for any possible entry decisions.

The airlines can enter markets by offering direct flights (for example, American Airlines offering a direct flight between Boston and Houston) and/or connecting flights (for example, American Airlines offering a flight between Boston and Houston with an intermediate stop at Miami). We formalise this as follows. It is useful to re-label each market t by keeping track of the endpoint cities. Specifically, we now denote the market between cities a and b by $\{a, b\}$. Given market $\{a, b\}$, let

$$G_{ab,f} = \begin{cases} 1 & \text{if airline } f \text{ offers direct flights between cities } a \text{ and } b, \\ 0 & \text{otherwise.} \end{cases}$$

Let $G_f \equiv (G_{ab,f} \ \forall \{a, b\} \in \mathcal{T})$ be the network of airline f where the *nodes* of the networks are the cities and the *links* of the network represent the markets served by airline f with direct flights. In the first stage, each airline f chooses its network G_f. In turn, we make this choice automatically determine which markets are served by airline f with connecting flights. In particular, if $G_{ah,f} = G_{hc,f} = 1$ and city h is one of airline f's hubs, then we assume that airline f also enters market $\{a, c\}$ by offering one-stop flights between cities a and c via h.

We assume that hub locations are exogenously determined before the starting of the game. This is because the transition from point-to-point to hub-and-spoke operations was a historical process started by the airlines after the U.S. Airline Deregulation Act of 1978 and quickly completed by the '90s, many years before the sample period of our empirical application. Once decided upon, hub locations were not altered in any major way by the airlines. For detailed studies on the transition to hub-and-spoke operations, see Caves et al. (1984), Kanafani and Ghobrial (1985), Morrison and Winston (1986), Levine (1987), Borenstein (1989; 1992), Butler and Houston (1989), Berry (1990), Brueckner et al. (1992), Evans and Kessides (1993), Brueckner and Spiller (1994), Oum et al. (1995), Berry et al. (1996), Nero (1999), Button, Forsyth, and Nijkamp (2000), and Reynolds-Feighan (2001). We also assume that the airlines can offer connecting flights with one stop only. Moreover, connections are possible at hubs only. This is because in our data we have almost no observations of connecting flights with more than one stop and connecting flights via non-hubs.

When entering markets, the airlines pay the fixed costs of building the physical, technological, and human infrastructures. Examples are the costs of salaries, insurance, scheduling coordination, computer reservation and revenue management system, and aircraft financing. The fixed costs also include the fees for ticket offices, baggage conveyor, gates, lounges, parking, and hangars at the airports. Further, as mentioned earlier, hub-and-spoke operations can increase the fixed costs due to the risk of congestion at hubs where many connections have to be wisely coordinated (for instance, see, Levine, 1987; Butler and Houston, 1989; Borenstein, 1992; Oum et al., 1995; Nero, 1999; Berry et al., 1996; Berry et al., 2019).

We specify the fixed costs sustained by airline f as:

$$\mathrm{FC}_f(G_f, \eta_f; \gamma) = \sum_{\{a,b\}\in\mathcal{T}} G_{ab,f}(\gamma_1 + \eta_{ab,f}) + \sum_{h\in\mathcal{H}_f} \gamma_{2,f}\Big(\sum_{\substack{a\,\in\,\mathcal{C}\\ a\,\neq\,h}} G_{ha,f}\Big)^2, \tag{2.7}$$

where $\mathcal{H}_f$ is the set of airline f's hubs, $\mathcal{C}$ is the set of cities, $\eta_f \equiv (\eta_{ab,f}\ \forall\{a,b\} \in \mathcal{T})$ is a vector of market-specific shocks that are observed by the airlines but are unobserved by the researcher, and $\gamma \equiv (\gamma_1, \gamma_{2,f}\ \forall f \in \mathcal{N})$ collects the parameters to be identified.[15] The fixed cost equation consists of two parts. First, there are market-specific contributions, $\gamma_1 + \eta_{ab,f}$, for each market $\{a,b\}$ served by airline f with direct flights. Second, there are quadratic terms, $\gamma_{2,f}(\sum_{\substack{a\,\in\,\mathcal{C}\\ a\,\neq\,h}} G_{ha,f})^2$, for each hub h of airline f, which account for the risk of congestion at hubs as discussed in the previous paragraph. In particular, $\sum_{\substack{a\,\in\,\mathcal{C}\\ a\,\neq\,h}} G_{ah,f}$ is the degree of hub h, i.e., the number of markets served out of hub h with direct flights (also called "spokes") by airline f.

Due to the quadratic terms in the fixed cost equation, the fixed costs sustained by airline f when serving a market out of hub h may depend on its decisions to serve other markets out of hub h. This gives rise to *spillover effects* in entry decisions across markets on the *fixed cost* side.

The assumption that the fixed cost shocks, $\eta \equiv (\eta_f\ \forall f \in \mathcal{N})$, are common knowledge among the airlines is deemed appropriate. In fact, in the airline industry, the fixed costs capture fairly standard balance sheet entries which pertain to the long-term side of the business and do not typically involve any industrial or technological secrets. Hence, it is plausible to suppose that the airlines are able to predict the competitors' fixed cost shocks reasonably well.[16]

Importantly, the fixed cost shocks are allowed to be correlated across markets and airlines. This is key because markets and airlines share endpoint cities. To establish point identification of the second-stage parameters, in Section 2.4.1 we will assume that the fixed cost shocks are uncorrelated with the second-stage shocks. In the same section, we will illustrate the pros and

[15] We could allow γ_1 to be firm-specific. We have not done so to maintain a parsimonious specification.

[16] The assumption that the fixed cost shocks are common knowledge is also imposed, for example, by Ciliberto et al. (2021) and Li et al. (2019).

cons of this restriction. Further, the fact that the airlines observe the fixed cost shocks when choosing their networks creates endogeneity issues which hamper the identification of the first-stage parameters. In Section 2.4.2, we will explain how to construct appropriate instruments, which are functions of some of the observed demand and marginal cost shifters. Apart from such instruments, the fixed cost shocks are allowed to be correlated with the observed demand and marginal cost shifters.

In the first stage, the airlines simultaneously choose the networks $G \equiv (G_f \ \forall f \in \mathcal{N})$ maximising the expected second-stage profits minus the fixed costs:

$$\text{profits of airline } f\text{:} \quad \mathbb{E}[\Pi_f(X^{\oplus}, W^{\oplus}, M, \xi^{\oplus}, \omega^{\oplus}, G; \theta)|X^{\oplus}, W^{\oplus}, M, \eta] - \text{FC}_f(G_f, \eta_f; \gamma), \quad (2.8)$$

where $\Pi_f(X^{\oplus}, W^{\oplus}, M, \xi^{\oplus}, \omega^{\oplus}, G; \theta)$ are the second-stage profits. Hereafter, we denote by $\mathcal{J}_t^{\oplus}$ the set of *all potential* products in market t, including the products not chosen for production. In turn, $X^{\oplus} \equiv (X_{j,t} \ \forall j \in \mathcal{J}_t^{\oplus} \ \forall t \in \mathcal{T})$, $W^{\oplus} \equiv (W_{j,t} \ \forall j \in \mathcal{J}_t^{\oplus} \ \forall t \in \mathcal{T})$, $\xi^{\oplus} \equiv (\xi_{j,t} \ \forall j \in \mathcal{J}_t^{\oplus} \ \forall t \in \mathcal{T})$, and $\omega^{\oplus} \equiv (\omega_{j,t} \ \forall j \in \mathcal{J}_t^{\oplus} \ \forall t \in \mathcal{T})$ are the vectors of observed demand shifters, observed marginal cost shifters, demand shocks, and marginal cost shocks of all potential products in every markets.[17] $M \equiv (M_t \ \forall t \in \mathcal{T})$ is the vector of market sizes. $\theta \equiv (\theta_d, \psi)$ is the vector of second-stage parameters. Note that the expectation of the second-stage profits is computed by integrating over the demand and marginal cost shocks, $(\xi^{\oplus}, \omega^{\oplus})$, conditional on the variables observed by the airlines in the first stage, $(X^{\oplus}, W^{\oplus}, M, \eta)$. We highlight that the second-stage profits depend on the networks formed by the airlines. In fact, the networks determine the competing firms, the offered products, the characteristics of the offered product, and the equilibrium prices in each market.

2.3.3 Equilibrium

The airlines solve the game by working backward from the second stage. First, they calculate the equilibrium profits under any possible networks, demand shocks, and marginal cost shocks. Then, they choose the networks maximising the expected value of those profits. A pure strategy subgame perfect Nash equilibrium consists of networks and price functions, $\{G^*, P_t^*(\xi_t^{\oplus}, \omega_t^{\oplus}, G) \ \forall t \in \mathcal{T}\}$, constituting a pure strategy Nash equilibrium in every subgame.

The existence and uniqueness of $\{P_t^*(\xi_t^{\oplus}, \omega_t^{\oplus}, G) \ \forall t \in \mathcal{T}\}$ is established by Nocke and Schutz (2018) for the case of multi-product Nested Logit, which is what we consider here.

We allow for multiple G^*. Multiple G^* are possible because the airlines compete at the entry stage through the second-stage pricing game. Our methodology does not require the existence

[17] Analogously, we define the market-specific vectors $X_t^{\oplus} \equiv (X_{j,t} \ \forall j \in \mathcal{J}_t^{\oplus})$, $W_t^{\oplus} \equiv (W_{j,t} \ \forall j \in \mathcal{J}_t^{\oplus})$, $s_t^{\oplus} \equiv (s_{j,t} \ \forall j \in \mathcal{J}_t^{\oplus})$, and $P_t^{\oplus} \equiv (P_{j,t} \ \forall j \in \mathcal{J}_t^{\oplus})$. We will use this notation also in Section 2.4.1.

of G^* for every possible parameterization and realization of the variables. In fact, as discussed in Section 2.4.2, we will use a collection of revealed-preference inequalities to bound the first-stage parameters, which are *implications* of (i.e., *necessary* conditions for) Nash equilibrium. If a particular parameterisation does not generate Nash equilibrium networks, then the revealed-preference inequalities *may* not be satisfied. In that case, this parameterization would not be included in the identified set. If no parameterization can satisfy the revealed-preference inequalities, then the identified set would be empty. We would conclude that the observed networks cannot be an equilibrium outcome under the model as specified, and so we might reject the model. Thus, our framework can be used even when nonexistence is possible. For further discussion on the existence of Nash equilibrium networks, see Appendix F.[18]

2.4 Identification

This section discusses identification of the vector of parameters, $(\theta_0, \gamma_0) \in \Theta \times \Gamma \subseteq \mathbb{R}^{\mathrm{K}} \times \mathbb{R}^{\mathrm{N}+1}$, where K is the dimension of θ_0, $\mathrm{N}+1$ is the dimension of γ_0, and the subscript "0" denotes the true parameter values.

2.4.1 Identification of the second-stage parameters

This section discusses identification of $\theta_0 \equiv (\theta_{d,0}, \psi_0) \in \Theta$. We follow the identification arguments of standard supply-demand models with differentiated products (Berry and Haile, 2014). Intuitively, the vector of demand parameters, $\theta_{d,0}$, is identified from the distribution of prices, sales, and product covariates. Once $\theta_{d,0}$ is identified, the markups are also identified from the F.O.C.s in (2.6). In turn, the marginal costs are identified as the difference between prices and markups. Finally, the variation in the identified marginal costs and product covariates identifies the vector of marginal cost parameters, ψ_0.

More precisely, there are two potential sources of endogeneity to be faced here. First, the list of products offered in the second stage is selected by the airlines in the first stage and may be correlated with the second-stage shocks. Second, the prices and within-group market shares may be correlated with the second-stage shocks because the latter are observed by the airlines when playing the second stage. We rule out the first source of endogeneity using a classic approach in empirical two-stage games: we assume that the second-stage shocks are mean independent of the airlines' information set in the first stage. The same assumption is imposed by Eizenberg (2014), Holmes (2011), Houde et al. (2017), Kuehn (2018), Rossetti (2018), and Wollmann (2018). We account for the second source of endogeneity by instrumenting the prices

[18]Note here the the revealed-preference inequalities are different, for example, from the identifying bounds in Ciliberto and Tamer (2009). The latter bounds are based on *necessary and sufficient* conditions for Nash equilibrium. Hence, they require the econometrician to explicitly deal with the potential case where the predicted set of Nash equilibria is empty.

and within-group market shares, as usual in supply-demand models. The point-identifying condition for θ_0 is formalised as follows:

Assumption 1. (*Exogeneity of the second-stage shocks*) For every market $t \in \mathcal{T}$ and product $j \in \mathcal{J}_t^{\oplus}$, $\mathbb{E}(\xi_{j,t}, \omega_{j,t} | X^{\oplus}, W^{\oplus}, M, \eta) = 0$. $\diamond$

Assumption 1 essentially states that the information owned by the airlines in the first stage does not help them to predict better the second-stage shocks. It is similar to the mean independence assumption in standard supply-demand models. However, there is a distinguishing aspect to notice, which relates to the two-stage structure of our game: here, we assume that the second-stage shocks are mean independent of the airlines' information set in the first stage, which includes also the covariates of the products not chosen for production and the fixed cost shocks.

Assumption 1 implies that $\mathbb{E}(\xi_{j,t}, \omega_{j,t} | G) = 0$ for every product j and market t, i.e., the second-stage shocks are mean independent of the list of products offered in the second stage. In turn, we can point identify θ_0 via the classic approach. Let $z_{j,t}(X_t^{\oplus}, W_t^{\oplus})$ be an $\mathrm{L} \times 1$ vector of instruments pertaining to product j in market t, where $\mathrm{L} \geq \mathrm{K}$. Given $\rho_{j,t} \equiv (\xi_{j,t}, \omega_{j,t})$, Assumption 2 allows us to write the following moment conditions:

$$\mathbb{E}(\rho_{j,t} \times z_{j,t,l}(X_t^{\oplus}, W_t^{\oplus}) | G) = 0 \quad \forall j \in \mathcal{J}_t^{\oplus}, \quad \forall t \in \mathcal{T}, \tag{2.9}$$

for every instrument $l = 1, \dots, \mathrm{L}$. Berry et al. (1995) show that we can uniquely express $\rho_{j,t}$ as a function of the product covariates and θ_0 ("BLP inversion"):

$$\rho_{j,t} = \tau_{j,t}(X_t^{\oplus}, W_t^{\oplus}, M_t, s_t^{\oplus}, P_t^{\oplus}, G; \theta_0) \quad \forall j \in \mathcal{J}_t^{\oplus}, \quad \forall t \in \mathcal{T}. \tag{2.10}$$

Therefore, we obtain:

$$\mathbb{E}(\tau_{j,t}(X_t^{\oplus}, W_t^{\oplus}, M_t, s_t^{\oplus}, P_t^{\oplus}, G; \theta_0) \times z_{j,t,l}(X_t^{\oplus}, W_t^{\oplus}) | G) = 0 \quad \forall j \in \mathcal{J}_t^{\oplus}, \quad \forall t \in \mathcal{T}, \tag{2.11}$$

for every instrument $l = 1, \dots, \mathrm{L}$. The above moment conditions depend only on variables that are observed by the researcher and provide point identification of θ_0. Inference on θ_0 can be conducted via GMM, as illustrated in Appendix G.1. As instruments for the prices and within-group market shares, we use functions of the observed demand shifters, as explained by Berry et al. (1995). For example, we consider the number of competing firms, the number of offered products, and the covariates of the competitors' products. In total, we have 13 instruments for each product.

We conclude the section with a discussion on the pros and cons of Assumption 1. Assumption 1 allows us to apply the classic identification approach to θ_0. Hence, it is key to feasibly nest a network formation step in a supply-demand framework, which is the main objective of our paper. Assumption 1 rules out correlation between the second-stage shocks and the fixed cost shocks. We view this as a reasonable simplification. In fact, the fixed cost shocks capture the residual fixed costs sustained to build the physical, technological, and human infrastructures. Hence, they pertain to the long-term side of the business. Instead, the second-stage shocks represent ticket restrictions and the quality of in-flight service. Hence, they generally relate to operational and short-term activities. Note here that the airline industry is different from the car industry, for instance, where producing a luxury car requires more up-front investment and greater fixed costs to create a single unit.[19] Assumption 1 rules out correlation between the second-stage shocks and the observed product characteristics, as standard in the literature on supply-demand models. Introducing some correlation between the second-stage shocks and the observed product characteristics can be an interesting, albeit difficult, extension for future work. Finally, to further rule out potential sources of correlations, recall that in the utility and marginal cost specifications we include hubbing variables, carrier fixed effects, and city fixed effects.

2.4.2 Identification of the first-stage parameters

This section discusses identification of $\gamma_0 \in \Gamma$. Identification of γ_0 is hampered by two main issues related to the discrete choice nature of the problem. First, there may be multiple Nash equilibrium networks. This is because the airlines compete at the entry stage through the second-stage pricing game. In turn, we are not able to write down a well-defined likelihood function. Second, even if one is willing to specify an equilibrium selection mechanism, it remains burdensome to construct the set of Nash equilibrium networks due to the large number of markets. In turn, we are not able to write down a tractable likelihood function that can be evaluated many times throughout an optimisation routine. Instead of focusing on Nash equilibrium networks, we bypass these two issues by considering implications of (i.e., necessary conditions for) Nash equilibrium, which are easier to handle from an econometric point of view. More precisely, we implement the revealed preference approach by Pakes (2010) and Pakes et al. (2015). This approach consists of two steps. First, we write down inequalities by simply predicting that the observed networks lead to higher profits than the profits would be were the airlines to deviate from the observed networks. These inequalities do not require that one solves for set of Nash equilibrium networks, do not rule out multiple equilibria, and do not restrict the selection mechanism used when there are multiple equilibria. Second, we get rid of the structural errors entering the inequalities by taking the expectation of the inequalities over markets and interacting them with instruments. The resulting moment inequalities characterise a non-sharp identified set that is a convex polytope, whose projections can be easily obtained

[19] Under the assumption that entry decisions are independent across markets, see Li et al. (2019) and Ciliberto et al. (2021) for possible ways to include correlation between the second-stage shocks and the fixed cost shocks.

by solving linear programming problems. In the remainder of the section, we formally illustrate the procedure.

We construct inequalities by considering *one-link* deviations from the observed networks. More precisely, let $G \equiv (G_f \; \forall f \in \mathcal{N})$ denote the observed networks. For each airline f and market $\{a, b\}$, if $G_{ab,f} = 1$, then we consider airline f deviating from G_f by not offering direct flights between cities a and b (hereafter, class of deviations "$(-ab)$"). Viceversa, if $G_{ab,f} = 0$, then we consider airline f deviating from G_f by offering direct flights between cities a and b (hereafter, class of deviations "$(+ab)$"). Such deviations should lead to lower profits and, hence, produce the following inequalities:

$$\begin{aligned} &\Delta\Pi_{(+ab),f} - \Delta\overline{\text{FC}}^{\top}_{(+ab),f}\gamma_0 + \eta_{ab,f} \geq 0 \quad \text{if } G_{ab,f} = 0, \\ &\Delta\Pi_{(-ab),f} - \Delta\overline{\text{FC}}^{\top}_{(-ab),f}\gamma_0 - \eta_{ab,f} \geq 0 \quad \text{if } G_{ab,f} = 1, \\ &\text{for each } \{a, b\} \in \mathcal{T} \text{ and } f \in \mathcal{N}. \end{aligned} \tag{2.12}$$

In (2.12), $\Delta\Pi_{(+ab),f}$ and $\Delta\Pi_{(-ab),f}$ denote the differences in the expected second-stage profits between the factual and counterfactual scenarios. They contain the second-stage parameters, which are assumed to be already identified. Similarly, $\Delta\overline{\text{FC}}^{\top}_{(+ab),f}\gamma_0$ and $\Delta\overline{\text{FC}}^{\top}_{(-ab),f}\gamma_0$ denote the differences in the systematic fixed costs between the factual and counterfactual scenarios. If $G_{ab,f} = 0$ (resp., $G_{ab,f} = 1$), then the fixed cost shock $\eta_{ab,f}$ should be added to (resp., subtracted from) the profit difference.

We highlight that, despite considering one-link deviations, the left-hand-side of the inequalities in (2.12) is not computed *as if* entry decisions were independent across markets. In fact, a one-link deviation creates a "domino effect" in the neighbour markets, due to the possibility for airline f to offer one-stop flights and the presence of spillover effects. This makes our method very different from the approaches which assume that entry decisions are independent across markets. As an example, consider the case where airline f deviates from G_f by now setting $G_{ab,f} = 1$. If city a is a hub for airline f and $G_{ac,f} = 1$ in the observed network, then airline f now competes also in market $\{b, c\}$ by offering one-stop flights between cities b and c via a. Similarly, if city b is a hub for airline f and $G_{bd,f} = 1$ in the observed network, then airline f now competes also in market $\{a, d\}$ by offering one-stop flights between cities a and d via b. Further, due to the presence of spillover effects on the demand side, the fact that now airline f offers direct flights between cities a and b could make it more attractive for customers to fly in any markets having cities a or b as origin. Also, due to the presence of spillover effects on the marginal cost side, the fact that now airline f offers direct flights between cities a and b could make it cheaper to offer flights having cities a or b as endpoints or intermediate stops. Finally, if cities a or b are hubs for airline f, then the fixed costs of offering direct flights from such hubs may increase due to congestion effects. Therefore, deviating to $G_{ab,f} = 1$ implies new equilibrium prices and fixed costs in market $\{a, b\}$ *and* in the markets that are in the neighbourhood of market $\{a, b\}$. All such effects are taken into account when computing the left-hand-side of the inequalities in (2.12). Appendix G.3 discusses in detail how we calculate $\Delta\Pi_{(+ab),f}$, $\Delta\overline{\text{FC}}_{(+ab),f}$, $\Delta\Pi_{(-ab),f}$, and $\Delta\overline{\text{FC}}_{(-ab),f}$.[20]

The inequalities in (2.12) cannot be used yet for identification because they contain the fixed cost shocks. In order to get of rid of the fixed cost shocks, we take the expectation of the inequalities in (2.12) over markets:

$$\begin{aligned}
&\mathbb{E}[\Delta\Pi_{(+ab),f} - \Delta\overline{\text{FC}}_{(+ab),f}^{\top}\gamma_0 + \eta_{ab,f}|G_{ab,f} = 0] \geq 0,\\
&\mathbb{E}[\Delta\Pi_{(-ab),f} - \Delta\overline{\text{FC}}_{(-ab),f}^{\top}\gamma_0 - \eta_{ab,f}|G_{ab,f} = 1] \geq 0,\\
&\text{for each } f \in \mathcal{N}.
\end{aligned} \tag{2.13}$$

If we could claim that $\mathbb{E}[\eta_{ab,f}|G_{ab,f} = 0]$ and $\mathbb{E}[\eta_{ab,f}|G_{ab,f} = 1]$ are equal to zero, then the moment inequalities in (2.15) could be used for identification. Unfortunately, such conditional expectations are not equal to zero because the fixed cost shocks represent structural components that are observed by the airlines when forming their networks.

We overcome this selection problem by introducing instruments, as discussed by Pakes et al. (2015). More precisely, suppose that for each airline f we have two positive variables, $Z_{(+ab),f}$ and $Z_{(-ab),f}$, such that:

$$\mathbb{E}[Z_{(+ab),f} \times \eta_{ab,f}|G_{ab,f} = 0] = 0 \quad \text{and} \quad \mathbb{E}[Z_{(-ab),f} \times \eta_{ab,f}|G_{ab,f} = 1] = 0. \tag{2.14}$$

We can interact these instruments with the moment inequalities in (2.15) and obtain:

$$\begin{aligned}
&\mathbb{E}[Z_{(+ab),f} \times (\Delta\Pi_{(+ab),f} - \Delta\overline{\text{FC}}_{(+ab),f}^{\top}\gamma_0 + \eta_{ab,f})|G_{ab,f} = 0] \geq 0,\\
&\mathbb{E}[Z_{(-ab),f} \times (\Delta\Pi_{(-ab),f} - \Delta\overline{\text{FC}}_{(-ab),f}^{\top}\gamma_0 - \eta_{ab,f})|G_{ab,f} = 1] \geq 0,\\
&\text{for each } f \in \mathcal{N}.
\end{aligned} \tag{2.15}$$

By the exogeneity restriction in (2.14), it holds that:

$$\begin{aligned}
&\mathbb{E}[Z_{(+ab),f} \times (\Delta\Pi_{(+ab),f} - \Delta\overline{\text{FC}}_{(+ab),f}^{\top}\gamma_0)|G_{ab,f} = 0] \geq 0,\\
&\mathbb{E}[Z_{(-ab),f} \times (\Delta\Pi_{(-ab),f} - \Delta\overline{\text{FC}}_{(-ab),f}^{\top}\gamma_0)|G_{ab,f} = 1] \geq 0,\\
&\text{for each } f \in \mathcal{N}.
\end{aligned} \tag{2.16}$$

The moment inequalities in (2.16) now depend only on variables that are observed by the researcher and, hence, can be exploited for identification. In particular, suppose that one

[20] For the same reasons, the left-hand-side of the inequalities in (2.12) is not computed as if each market was run by a local manager taking independent entry decisions as assumed in Aguirregabiria and Ho (2012).

disposes of R instruments satisfying the exogeneity restriction in (2.14). Then, one can use these instruments to obtain R moment inequalities which characterise a non-sharp identified set for γ_0. Further, if the instruments offer sufficient variation in profits relative to the fixed cost shocks, then the identified set is bounded. Note also that the moment inequalities in (2.16) are linear in γ_0. Therefore, the identified set is a convex polytope.

To construct our first-stage instruments, we follow the approach implemented by Wollmann (2018) in an empirical study of the commercial truck production process. Wollmann (2018) suggests to think about $Z_{(+ab),f}$ as being a function of the information set of firm f in the first stage, taking value 1 if offering direct flights between cities a and b is, *on average*, unoptimal for firm f regardless of $\eta_{ab,f}$, and 0 otherwise. Similarly, $Z_{(-ab),f}$ is a function of the information set of airline f in the first stage, taking value 1 if offering direct flights between cities a and b is, *on average*, optimal for airline f regardless of $\eta_{ab,f}$, and 0 otherwise. For example, we set $Z_{(-ab),f} = 1$ if market $\{a, b\}$ has a very large size and cities a or b are hubs for airline f. In fact, this indicator isolates markets where airline f will tend to always offer direct flights, plausibly unrelated to the fixed cost shocks, due to the expected very high profitability. Viceversa, we set $Z_{(+ab),f} = 1$ if market $\{a, b\}$ has a very small size and cities a and b are hubs for the competitors and not for airline f. In fact, this indicator isolates markets where airline f will tend to never offer direct flights, plausibly unrelated to the fixed cost shocks, due to the expected very low profitability. Along these lines, we construct 5 instruments per each airline which are discussed in Appendix G.4.

Instead of introducing instruments, an alternative approach for getting rid of the fixed cost shocks from the inequalities in (2.12) consists of introducing support restrictions on the fixed cost shocks. For example, in an empirical study of the personal computer industry, Eizenberg (2014) assumes that the fixed cost shocks have a bounded support which is contained within the support of the expected change in the second-stage profits resulting from one-product deviations at a time. We have not pursued this approach because, by construction, it characterises an unbounded identified set when there are spillover effects in entry decisions on the fixed cost side, as explained by Eizenberg (2014) in Appendix A.3.

The identified set that we characterise is not sharp. In principle, one could construct revealed-preference inequalities based on "richer" classes of deviations, for example when an airline adds or deletes two or more links simultaneously. If one was able to find appropriate instruments, the resulting additional moment inequalities would sharpen the identified set. We have not considered simultaneous-link deviations for two reasons. First, finding appropriate instruments for simultaneous-link deviations is challenging. Second, in the empirical application, our strategy delivers sufficiently informative bounds for the first-stage parameters.

We summarise the above discussion with the following assumption:

Assumption 2. (*First-stage instruments*) For each airline f, there exists variables $Z^{r+}_{(+ab),f}$ for $r_+ = 1,\dots,\mathrm{R}_+$ and $Z^{r-}_{(-ab),f}$ for $r_- = 1,\dots,\mathrm{R}_-$ such that:

$$\begin{aligned}
&\mathbb{E}[Z^{r+}_{(+ab),f} \times \eta_{ab,f} | G_{ab,f} = 0] = 0 \quad \text{for } r_+ = 1,\dots,\mathrm{R}_+,\\
&\mathbb{E}[Z^{r-}_{(-ab),f} \times \eta_{ab,f} | G_{ab,f} = 1] = 0 \quad \text{for } r_- = 1,\dots,\mathrm{R}_-.
\end{aligned}$$

◇

In turn, the identified set for γ_0 is:

$$\begin{aligned}
\Gamma_I \equiv \Big\{\gamma \in \Gamma : &\mathbb{E}[Z_{(+ab),f} \times (\Delta\Pi_{(+ab),f} - \Delta\overline{\mathrm{FC}}^\top_{(+ab),f}\gamma_0) | G_{ab,f} = 0] \geq 0,\\
&\mathbb{E}[Z_{(-ab),f} \times (\Delta\Pi_{(-ab),f} - \Delta\overline{\mathrm{FC}}^\top_{(-ab),f}\gamma_0) | G_{ab,f} = 1] \geq 0,\\
&\text{for } r_+ = 1,\dots,\mathrm{R}_+,\ r_- = 1,\dots,\mathrm{R}_-,\ f \in \mathcal{N}\Big\}.
\end{aligned} \tag{2.17}$$

In Appendix G.2, we discuss how to conduct inference on Γ_I.

2.5 Empirical application

2.5.1 Data

We use data from the Airline Origin and Destination Service (hereafter, DB1D) which consists of a 10% random sample of all the tickets issued in the United States during the second quarter of 2011. By then, the merger between United Airlines and Continental Airlines had been completed and American Airlines and US Airways had not announced their intention to merge yet. We restrict the sample to the flights operated between the 85 largest metropolitan statistical areas (MSAs) in the United States. We refer to MSAs as cities throughout the section. Further, if a city has more than one airport (such as New York, Chicago, and Los Angeles), we combine its airports into one.[21] If an airport within a city serves as a hub for a given airline, then that city is a hub for the airline.[22] The major carriers in the sample are United Airlines (UA), Delta Airlines (DL), American Airlines (AA), US Airways (US), and Southwest Airlines (WN). All the other carriers in the sample are put either in a group called "Low Cost Carriers" (LCC), or in a group called "Other". These carriers are treated as fringe competitors, differing only in whether or not they can be classified as low cost. Also, to enhance computational tractability, we do not consider their fixed costs when estimating

[21] See Footnote 7.

[22] For instance, Dallas/Fort Worth serves as a hub for American Airlines whereas Dallas Love Field does not. Given that we combine both airports into one, the resulting city (Dallas) is a hub for American Airlines.

the first stage parameters and we assume that their networks are exogenously pre-determined before the starting of the game. The legacy carriers use hub-and-spoke operations. Southwest Airlines does not rely on a pure hub-and-spoke business model, but rather on a hybrid system in which some airports are "focus cities" offering some, but not all, of the services generally found at hubs. When estimating our model, we treat "focus cities" as hubs.[23]

We delete tickets with multiple operating carriers or multiple ticketing carriers. Also, we delete tickets with different inbound and outbound itineraries. Further, we delete tickets that are not round-trip. Lastly, to be consistent with our model, we delete connecting tickets via cities that are not hubs. Note that we observe very few of these tickets. As discussed in Section 2.3, we consider tickets featuring the same airline-itinerary combination but different fares as the same product. We compute the corresponding price as follows. First, we delete tickets with fares in the highest and lowest percentiles and tickets with fares below $25. Then, we construct the weighted average price over the remaining fares.

We allow the marginal cost parameters to differ between short-haul and long-haul flights, which are defined as flights covering up to 1,500 miles and flights covering more than 1,500 miles, respectively. As anticipated in Section 2.3, for each market $t \in \mathcal{T}$ and product $j \in \mathcal{J}_t^{\oplus}$, $X_{j,t}$ collects the number of stops ("Stops"), the maximum number of direct flights offered at the itinerary's endpoints by the same carrier offering itinerary j ("Connections"), the distance flown ("Distance"), and its squared value ("Distance2"). Similarly, $W_{j,t}$ collects the number of stops ("Stops Short", "Stops Long"), the average number of cities that are reachable from the endpoints and intermediate stops of itinerary j with the same carrier offering itinerary j ("Presence Short", "Presence Long"), and the distance flown ("Distance Short", "Distance Long"). We include in the demand and supply models firm and city fixed effects in order to capture brand preferences and unobserved city-specific features. Lastly, we compute market sizes using data from the US Census Bureau on MSA population. In particular, we calculate the size of each market $t \in \mathcal{T}$, M_t, as the geometric mean of the populations at the market's endpoints.

Table 2.1 provides some summary statistics. In panel (a), we see that whereas only around 15% of flights are direct, they account for around 85% of passengers. We can also see the importance of hubs: 57% of passengers travel through, from, or to hubs and 83% of all flights start, land, or connect at a hub. Further, we see in panel c) that airlines serve on average almost 50 direct flights out of hub airports (measured by the "Degree" variable), compared to around 7 for non-hub airports. Looking at panel e), we see that we have on average 5.56 products per market and on average 4.62 hub products.

[23] See Table H.0.2 in Appendix H for the list of hubs.

Table 2.1: Summary statistics

(a) Sizes		
Number of firms	7	
Number of products	17,481	
Number of markets	3,146	
Fraction of direct flights	0.14	
Fraction of hub itineraries	0.83	
Fraction of direct passengers	0.85	
Fraction of passengers in hub markets	0.57	
Fraction of markets served	0.93	
(b) Passengers by airline (1 million)		
Total	25.33	
American	3.15	
Delta	4.85	
United	3.81	
US Airways	2.21	
Southwest	6.00	
Low Cost	4.10	
Other	1.21	
(c) Network statistics	**Mean**	**St.Dev**
Degree (Hub)	49.86	13.03
Density (Hub)	0.61	0.16
Clustering (Hub)	0.24	0.14
Degree (Non-hub)	7.21	7.72
Density (Non-hub)	0.09	0.09
Clustering (Non-hub)	0.80	0.33
(d) Demand and marginal cost variables	**Mean**	**St.Dev**
Price (100 USD)	4.32	1.20
Stops	0.86	0.34
Connections (100)	0.20	0.19
Presence (100)	0.56	0.15
Distance (100 km)	1.44	0.68
Product share	4.6083e-04	1.4784e-03
Market size (1 million)	2.55	1.85
(e) Market-level statistics	**Mean**	**St.Dev**
Number of firms	3.59	1.81
Number of products	5.56	4.43
Number of direct flights	0.75	1.20
Number of hub itineraries	4.62	3.43
Number of passengers (1,000)	8.05	24.43
Number of direct passengers (1,000)	6.82	23.98
Number of passengers in hub markets (1,000)	4.60	15.39

Note:
Hub itineraries are itineraries where at least one of the endpoints or intermediate stop is a hub. Hub markets are markets where at least one of the endpoints is a hub. The share of a product is computed as the total number of passengers buying that product divided by the market size, times 10 because we have a 10% random sample. The degree of a (non-)hub is the number of markets served with direct flights out of the (non-)hub by an airline. The density of a (non-)hub is the ratio between the number of markets served with direct flights out of the (non-)hub by an airline and the total number of potential markets out of the (non-)hub. The clustering coefficient of a (non-)hub is the ratio between the number of closed triplets including the (non)-hub served by an airline and the total number of potential triplets including the (non)-hub.

2.5.2 Results from the second stage

The second-stage results are in Table 2.2. We find significant spillover effects in entry decisions across markets on the demand side. Specifically, passengers benefit from having a large number of direct flights offered by an airline at the itinerary's endpoints (Connections). That is, dense hubs increase passengers utility, all the rest being constant. This effect captures the value of frequent flier programs. In fact, the larger the number of destinations for which consumers can redeem frequent flier miles, the higher the value of such loyalty programs. Additionally, an airline that flies to many cities is likely to have more convenient parking and gate access and provide better services. The price coefficient is negative. It lies between the price coefficients of the two consumer types considered by Berry and Jia (2010) and within the ballpark of what other contributions have found. Consumer utility is an inverted U-shaped function of the distance flown. This means that, as distance increases, air travel becomes more attractive relative to the outside option. However, as distance increases further, travel becomes less pleasant and demand starts to decrease. In line with the literature, passengers exhibit a strong disutility for connecting flights. Lastly, we estimate the nesting parameter, λ, to be around 0.6. Recall that, as λ approaches one, the Nested Logit model reduces to the standard Logit model. Therefore, we can conclude that there is substitution between the inside goods and the outside option.

We find significant spillover effects in entry decisions across markets also on the marginal cost side. Specifically, the marginal cost of an itinerary decreases with the average number of cities that an airline allows to reach from the endpoints and intermediate stops (Presence). This effects captures the impact of economies of density on the marginal costs, i.e., the fact that more densely travel segments tend to have lower unit costs due to engineering reasons. In particular, the larger the number of final destinations consumers can reach, the more the opportunities for an airline to pool passengers from several itineraries into the same planes, and the more an airline can efficiently use large aircrafts which typically have lower unit costs. This mechanism implies that hub-and-spoke operations provide marginal cost savings, all the rest being constant. The impact of the variable Presence is more pronounced for long-haul flights, as the efficiency of large planes is especially evident in long routes. Further, the number of stops in long-haul flights reduces their marginal cost. Again, this suggests that connecting flights are less expensive to provide for long routes, by virtue of economies of density. The number of stops does not significantly reduce the marginal cost of short-haul flights. This may be because the marginal cost savings induced by economies of density are balanced by the extra take-off and landing, which uses a lot of fuel. The marginal cost of both long-haul and short-haul flights increases with the distance flown. This is because the longer the distance, the more fuel is needed to cover it. Lastly, as expected, Southwest Airline, Low Cost, and Other have lower marginal costs than the legacy carriers.

Table 2.2: Second-stage estimates

Utility			**Marginal Cost**		
	Coefficient	SE		Coefficient	SE
Mean utility			**Short-haul flights**		
Intercept	-5.598	(0.262)	Intercept	3.118	(0.090)
Price	-0.587	(0.066)	Stops	0.031	(0.028)
Stops	-1.794	(0.066)	Distance	0.474	(0.037)
Connections	0.868	(0.032)	Presence	-1.245	(0.136)
Distance	0.289	(0.084)	**Long-haul flights**		
Distance2	-0.093	(0.095)	Intercept	3.703	(0.114)
Nesting Parameter (λ)	0.623	(0.025)	Stops	-0.189	(0.041)
			Distance	0.667	(0.032)
			Presence	-2.016	(0.145)
Carrier FEs			**Carrier FEs**		
DL	-0.168	(0.018)	DL	0.082	(0.035)
UA	-0.387	(0.025)	UA	0.050	(0.032)
US	0.142	(0.025)	US	0.079	(0.032)
WN	-0.519	(0.032)	WN	-0.363	(0.029)
LCC	-0.348	(0.032)	LCC	-1.509	(0.055)
Other	-0.074	(0.056)	Other	-1.398	(0.049)
Statistics					
J-statistic	15.627				
Number of observations	17,481				
Price Elasticity	-3.780				
Aggregate Easticity	-2.100				
Connection semi-elasticity	0.880				

Note:
Prices are divided by USD 100. Connections and Presence are divided by 100. City fixed effects are included. The number of over-identifying restrictions is 11.

The bottom part of Table 2.2 reports some elasticity estimates. In particular, the price elasticity is the average estimated price elasticity across products. The aggregate elasticity is the percentage change in the inside product share when all prices rise by 1%. The connection semi-elasticity is the change in the inside product share if all direct flights became one-stop flights, while holding the other characteristics fixed. The price elasticity is slightly higher than in other contributions in the literature. This may be due to the fact that our model does not capture sufficient consumer heterogeneity in price sensitivity. The connection semi-elasticity is higher than in Berry and Jia (2010). This is in line with the trend towards increasingly strong preferences for direct flights found in their paper.

Table 2.3: Profits by firms

	Profits (100k)	Price	Marginal cost	Markup	Lerner Index
AA					
All	1.78	453.36	335.20	118.16	0.28
Direct	13.77	402.37	277.42	124.94	0.32
One-stop	0.39	459.26	341.89	117.38	0.27
Direct, hub endpoint	15.06	402.75	276.66	126.09	0.33
Direct, non-hub endpoints	2.00	398.87	284.48	114.40	0.30
DL					
All	1.41	436.45	310.40	126.05	0.31
Direct	12.31	463.26	321.03	142.23	0.33
One-stop	0.33	433.80	309.35	124.45	0.31
Direct, hub endpoint	13.49	482.67	336.83	145.84	0.32
Direct, non-hub endpoints	4.47	334.75	216.44	118.31	0.38
UA					
All	1.25	445.56	328.43	117.13	0.28
Direct	9.17	458.50	334.97	123.53	0.29
One-stop	0.20	443.85	327.56	116.28	0.28
Direct, hub endpoint	11.03	456.82	332.24	124.58	0.29
Direct, non-hub endpoints	2.17	464.88	345.33	119.55	0.29
US					
All	1.30	453.43	336.77	116.67	0.27
Direct	8.99	407.34	275.17	132.17	0.35
One-stop	0.35	459.10	344.34	114.76	0.26
Direct, hub endpoint	10.42	418.96	282.96	136.00	0.35
Direct, non-hub endpoints	3.95	366.22	247.58	118.64	0.36
WN					
All	2.79	419.43	299.51	119.92	0.31
Direct	12.09	365.14	237.09	128.05	0.38
One-stop	0.23	434.40	316.73	117.67	0.29
Direct, hub endpoint	16.49	362.34	233.95	128.39	0.38
Direct, non-hub endpoints	8.88	367.19	239.39	127.80	0.38

Table 2.3 reports firm-level profits, prices, marginal costs, and markups averaged over different products. For each airline, the first row contains the average across all products, the second line across direct flights, and the third line across one-stop flights. The fourth and fifth rows contain the average across direct flights where at least one of the endpoints is a hub (hub markets) and where no endpoint is a hub (non-hub markets), respectively. We can see that

the airlines charge a higher markup on direct flights compared to one-stop flights, which is in line with the fact that consumers value direct flights more. The legacy carriers charge a higher markup on direct flights in hub markets, compared to non-hub markets, suggesting the presence of a hub premium. Charging a high markup on hub markets may be due to high market power at hubs, or to high fixed costs from managing hubs. Whereas American Airlines, US Airways, and Southwest Airlines have substantially lower marginal costs on nonstop flights, it is the opposite for Delta and United Airlines. The marginal cost of Southwest Airlines is lower than the marginal costs of the legacy carriers. For direct flights, the difference is quite substantial. For one-stop flights, Southwest Airlines's advantage is small. The last finding is line with the fact that Southwest Airlines uses focus cities, rather than hubs. Hence, the marginal cost savings from offering connecting flights may be less pronounced since not all features of traditional hubs are exploited.

2.5.3 Results from the first stage

Table 2.4: Projections of the estimated identified set

Variable	Lower bound	Upper bound
Intercept	672,624	1,047,511
Congestion costs		
AA	19,216	28,024
DL	12,824	21,776
UA	8,731	16,586
US	27,191	39,044
WL	17,346	30,967

Note:
Entry costs are in $.

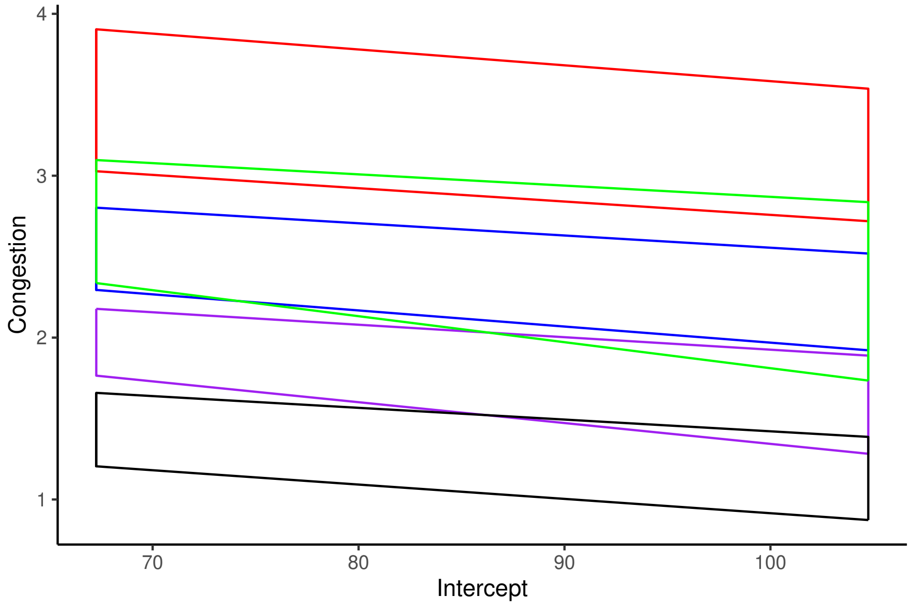
Table 2.4 reports the projections of the estimated identified set.[24] Estimates are in dollars. We see that, absent congestion costs, offering direct service between two endpoints costs between $\$672,624$ and $\$1,047,511$. To interpret the congestion effect, recall from our specification of the fixed costs (2.7) that the congestion costs of managing hubs are quadratic in the number of spokes of each hub. For example, American Airlines has a total of 221 spokes, resulting in congestion costs between \$214 and \$312 million.[25] Their base-line entry costs amount to between \$157 and \$245 million. The base-line entry cost is just the number of segments served multiplied by the Intercept term. Due to congestion costs, the fixed costs sustained by an airline when serving a market out of a hub depends on its decisions to serve other markets out of the hub. This gives rise to spillover effects in entry decisions across markets on the fixed

[24]We have not implemented the inference methodology yet.

[25]To compute congestion costs, we need to count the number of spokes at each hub, square that number, sum it, and then multiply by the estimated congestion cost. For American, this computation yields, for the lower bound, $19,216 * (68^2 + 29^2 + 26^2 + 39^2 + 59^2)$

cost side. Congestion costs differ substantially across the airlines. US Airways faces larger congestion costs than the other legacy carriers. United Airlines faces lower congestion costs than the other legacy carriers. The heterogeneity in congestion costs is also highlighted by Figure 2.2 The Figure shows the projections of the estimated identified sets for the 5 airlines. We see that while some sets overlap, there are substantial differences between airlines.

Figure 2.2: Projections of the estimated identified set



Colors: American (blue), Delta (purple), United (black), US Airways (red), and Southwest (green)

Table 2.5: Predicted entry probabilities

Firm	Data	Predicted
American	6.89 %	9.74 %
Delta	12.34 %	11.16 %
United	12.96 %	8.41 %
US Airways	6.95 %	11.12 %
Southwest	18.59 %	32.92 %

In Table 2.5, we report the predicted entry probabilities. To compute the predicted entry probabilities, we construct a grid of admissible parameter values by taking 100,000 draws from the convex polytope defined by our moment inequalities via Gibbs sampling. Then, for each airline, we implement the following procedure. For every parameter value and market, first, we compute the marginal profit from serving the market with direct flights without considering the fixed costs shock; second, we save one if the marginal profit is positive, and zero otherwise. For each parameter value, we sum all the ones and divide the result by the number of markets. Finally, we take the midpoint of these numbers across parameter values. This is the predicted entry probability. Overall, we predict entry patterns reasonably well. The largest discrepancy occurs for Southwest Airlines, which may be due to the fact that Southwest Airlines relies on focus cities, rather than hubs, whose peculiarities are not entirely captured by our framework.

2.5.4 Counterfactuals

This section studies the impact on firm and market outcomes of a merger between two of the four legacy carriers in our sample, American Airlines and US Airways. These two firms did in fact merge in 2013. They first expressed interest to merge in January 2012 and officially announced their plans to merge in February 2013. At the time they expressed interest to merge, American Airlines' holding company (AMR Corporation) was in Chapter 11 bankruptcy.[26] The Department of Justice (DoJ), along with several state attorney generals, sought to block the merger, concerned that the merger would have substantially lessen competition and hurt consumers. In 2013, a settlement was reached in which the merging parties pledged to give up landing slots or gates at 7 major airports and "*to maintain hubs in Charlotte, New York (Kennedy), Los Angeles, Miami, Chicago (O'Hare), Philadelphia, and Phoenix consistent with historical operations for a period of three years*".[27] Below, we refer to such settlement as the 2013 settlement. According to articles from the time the merger was announced, the parties expected the merger to make the new entity the largest airline in the world in terms of passenger numbers, and annual cost savings of around $1 billion per year.[28] Also, the merger was seen by analysts as an opportunity for American Airlines to expand its footprint in markets along the East Coast, where US Airways had a strong presence.[29]. The merger was the last in a series of large airline mergers and reduced the number of legacy carriers to 4 (Delta Airlines, United Airlines, Southwest Airlines, and the new American Airlines).

We simulate two counterfactual events. In the first event, we assume that American Airlines and US Airways merge. In the second event, we assume that American Airlines goes bankrupt and disappears. We consider the bankruptcy event because, at the time of the merger's announcement, American Airlines was indeed in Chapter 11 bankruptcy, which raises the question of what would have happened had it just disappeared from the market. Note that, throughout the first decade of the new millennium, all the legacy carriers filed for bankruptcy at some point, but were allowed to re-structure or merge in order to recover from financial distress. The debate on the appropriateness of aid to airlines that struggle financially has again become a matter of public concern during the Covid-19 pandemic. Thus, the analysis of the bankruptcy event aims to offer insights on this topic.

[26]Recall that we use data from the second quarter of 2011. This is before the two parties expressed interest to merge and corresponds to the last quarter before AMR filed for Chapter 11 bankruptcy.

[27]https://www.justice.gov/opa/pr/justice-department-requires-us-airways-and-american-airlines-divest-facilities-seven-key, https://americanairlines.gcs-web.com/news-releases/news-release-details/amr-corporation-and-us-airways-announce-settlement-us-department

[28]https://www.reuters.com/article/uk-americanairlines-merger-idUSLNE91D02020130214

[29]https://money.cnn.com/2013/02/14/news/companies/us-airways-american-airlines-merger/index.html

Set-up

When evaluating the merger event, we compare 5 scenarios:

1. *Networks fixed - Base case.* After the merger, the networks remain at the pre-merger levels. The firms maintain the pre-merger products and dummies. If the merging firms offer the same itinerary, then the two products are kept as separate. The firms play the simultaneous pricing game described in Section 2.3.1 and new equilibrium prices arise. In particular, the merging firms choose the prices maximizing their joint profits, i.e., they behave as if they colluded.

2. *Networks fixed - Best case.* After the merger, the networks remain at the pre-merger levels. All the firms, except American Airlines and US Airways, maintain the pre-merger products and dummies. The merging firms maintain the pre-merger products, but update some of their covariates. In particular, the products of the merging firms inherit the best firm dummies. If the merging firms offer the same itinerary, then the two products are kept as separate. However, differently from the previous scenario, now the two products inherit the most favourable observed demand and marginal cost shifters. For example, on the demand side, the estimated coefficient of the variable Connections is positive. Hence, the two products get the highest value of Connections between what American Airlines and US Airways had before merging. After such rearrangements, the firms play the simultaneous pricing game described in Section 2.3.1 and new equilibrium prices arise. As in the previous scenario, the merging firms choose the prices maximizing their joint profits, i.e., they behave as if they colluded.

3. *Networks fixed - Updated case.* After the merger, the networks of all the firms, except American Airlines and US Airways, remain at the pre-merger levels. We treat the merged entity as a new firm and assign it the network resulting from merging the pre-merger networks of American Airlines and US Airways. The products of the merged entity and their covariates are constructed from the merged network. The merged entity takes on the most favourable dummies of the merging firms. The demand and marginal cost shocks of the products of the merged entity stay the same as pre-merger, except in markets where both firms were present. There, we use the mean of the pre-merger errors. After such rearrangements, the merged entity and the other firms play the simultaneous pricing game described in Section 2.3.1 and new equilibrium prices arise.

4. *Networks vary - No remedies.* After the merger, we treat the merged entity as a new firm and we let the firms play the entire two-stage game described in Section 2.3. New equilibrium networks and prices arise. More details on how the firms reoptimise networks and prices are

in Section 2.5.4.

5. *Networks vary - With remedies.* After the merger, we treat the merged entity as a new firm and we let the firms play the entire two-stage game described in Section 2.3. New equilibrium networks and prices arise. However, differently from the previous scenario, now we take into account some of the remedies imposed to the merged entity by the 2013 settlement. In particular, recall that the 2013 settlement invited the merged entity "*to maintain hubs in Charlotte, New York (Kennedy), Los Angeles, Miami, Chicago (O'Hare), Philadelphia, and Phoenix consistent with historical operations for a period of three years*". We incorporate these remedies as binding constraints and force the merged entity to keep serving all the markets served pre-merger by the merging firms if one or both endpoints are at one of the hubs mentioned in the 2013 settlement. Note that this scenario differs from the "Networks fixed - Updated case" scenario because, first, the competitors of the merged entity are allowed to reoptimise their networks; second, the merged entity is allowed to enter all markets and exit those markets whose endpoints were not subject to the remedy.

The first three scenarios assume that the airlines do not reoptimise their networks after the merger, as standard in the literature. In turn, the analyst should make ad-hoc assumptions on how the products of the merged entity adjust after the merger, which opens infinite possibilities. The first three scenarios are just some examples and do not obviously exhaust all potential cases, with consequent risk of misspecification. The fourth and fifth scenarios consider the entire two-stage game and allow the firms to reoptimise their prices *and* networks after the merger, by leveraging on our methodology. In fact, after the merger, it is plausible to believe that the merged firm and its competitors will react not only by adjusting their prices, but also by repositioning in markets. For example, after the merger, there might room in some markets for accommodating other entrants. Further, the merger could generate marginal cost savings for the merged firm, by virtue of economies of density triggered by hub-and-spoke operations, which may favour its entry in new markets. The merger could also increase the market power of the merged firm, by disposing of a larger network that enhances consumer willingness to pay. At the same time, the merger might increase the total fixed costs of the merged firm, due to congestion effects at hubs, which may force it to dismiss some operations. All such synergies across markets are taken into account by our procedure.

When evaluating the bankruptcy and disappearance of American Airlines, we consider the following two scenarios:

1. *Networks Fixed.* After the disappearance of American Airlines, the networks of the other firms remain at the pre-merger levels. The firms play the simultaneous pricing game described in Section 2.3.1 and new equilibrium prices arise.

2. *Networks Vary.* After the disappearance of American Airlines, we let the other firms play the entire two-stage game described in Section 2.3. New equilibrium networks and prices arise.

The first scenario assumes that the airlines do not reoptimise their networks after the disappearance of American Airlines, as standard in the literature. The second scenario considers the entire two-stage game and allow the firms to reoptimise their prices *and* networks after the disappearance of American Airlines, by leveraging on our methodology. In fact, after the disappearance of American Airlines, it is plausible to believe that its competitors will react not only by adjusting their prices, but also by repositioning in markets. For example, we expect that the disappearance of American Airlines makes it more attractive for other firms to enter markets previously served by American Airlines, in turn alleviating the negative effects of the bankruptcy on consumer surplus. At the same time, replacing the hub-and-spoke operations of the disappearing airline may be infeasible due to the large fixed costs. It is then likely that consumers living in the hub cities of American Airlines will be overall worse off as they will no longer be able to benefit from the services previously offered by American Airlines.

Description of the counterfactual algorithm

In this section, we describe the algorithm implemented to reach a new equilibrium when the airlines reoptimise their prices and networks, in the "Networks vary - No remedies" scenario. In the "Networks vary - With remedies' scenario, we follow the same procedure, but we force the merged entity not to exit all the markets served pre-merger by the merging firms if one or both endpoints are at one of the hubs mentioned in the 2013 settlement.

Recall that there can be multiple equilibrium networks. Hence, in principle, it would be desirable to enumerate all possible equilibrium networks that may arise, as in Eizenberg (2014). However, doing so is infeasible in our setting due to the large number of markets. To make the problem tractable, we build a best-response learning algorithm based on Lee and Pakes (2009) and Wollmann (2018). In particular, we order markets and firms according to some criteria. For a given value of the parameters, firms iteratively best-respond to one another with respect to entry and price decisions sequentially over markets until they reach convergence. We are currently experimenting many different orders of markets and firms in order to obtain a *distribution* of equilibria. For the moment, the results reported below correspond to one specific order of markets and firms. We repeat the procedure for 25 draws of parameter values from the estimated identified set and report the minimum and maximum changes in the networks and market outcomes across such parameter values. In what follows, we provide more details on the steps of the algorithm, for a given value of the parameters and for a given order of firms and markets.

1. We rank markets according to: whether at least one of the market's endpoints are hubs for the merging airlines; whether the merging airlines served the market; the number of markets the merging airlines served out of the endpoints with direct flights; the market size.

2. We rank firms in the following order: American Airlines; Delta Airlines; United Airlines; US Airways; Southwest. When simulating the merger event, we initially assign to the merged entity the network resulting from merging the pre-merger networks of American Airlines and US Airways and let merged entity move first. We also assume that the cities in which either American Airlines or US Airways had a hub prior to the merger will continue to serve as hubs. This means that the merged entity will entertain hubs in Dallas, Chicago, Charlotte, Philadelphia, New York City, Washington DC, Phoenix, Miami, and Los Angeles. Further, the merged entity takes on the most favourable firm dummies of the merging firms.

3. For a given firm in a given market, we let the firm play its best response, holding the firm's network outside of the considered market and the rival networks and prices fixed at the level reached in the previous iteration. In order to find the best response of the firm, we compute the firm's total second-stage profits when serving the market with direct flights, the total second-stage profits when not serving the market with direct flights, and take the difference. Note here that we let the firm to best respond with respect to prices both in the market under consideration and in the neighbour markets due to spillover effects and the possibility of offering one-stop flights. Hence, the simulation is not conducted *as if* entry decisions were independent across markets. We also compute the total fixed costs when the firm serves the market with direct flights and when it does not, and take the difference. If the second-stage profit difference is larger (smaller) than the fixed cost difference, then the best response of the firm is to (not) serve the market with direct flights. We update the network of the firm according to the best response and move to the next firm. We cycle through firms in a given market until no firm wishes to deviate.

4. We cycle through the markets and check how many entry decisions have changed. If that number is larger than some tolerance criterium, we repeat the entire procedure. Once the number of changed decisions is below the criterium, we stop the procedure.

Note that, at the rest point of the procedure described above, the necessary conditions that are used in the estimation of first-stage parameters hold. Hence, the procedure provides an equilibrium that is is internally consistent with our model. Note also that computational costs prohibit to consider all possible entry deviations by each firm, although we believe this results in no meaningful loss of generality. In this respect, the equilibrium reached by the above procedure is a Nash equilibrium within the classes of entry deviations considered. For example, we impose that no firm considers adding/deleting direct flights in more than one market at a time. After extensive experimentation, we concluded that no firm would best respond with more changes than that. However, recall that the airlines offer also one-stop flights and, thus, the total number of product changes at each iteration can be greater than one. Further, we allow the airlines add/delete direct flights only in the hub markets of American Airlines and US Airways. These markets represent around 20% of all segments in our sample and are presumably those where the DoJ would be most worried about potential anti-competitive effects of the merger.

Additional details on how the latent variables are imputed

To perform counterfactuals, we need a measure of the fixed cost shocks. Different approaches have been taken in the literature. For example, Wollmann (2018) draws the fixed cost shocks from a normal distribution with zero mean and variance equal to a fraction of the variance of the systematic fixed costs. Kuehn (2018) finds, for each market, the range of realisations of the fixed cost shock generating either entry or exit and takes the midpoint. We use a procedure that is similar in spirit to Kuehn (2018). In particular, when we observe airline f serving market $\{a, b\}$ with direct flights, we subsume that this choice must be profitable, giving us an upper bound for $\eta_{ab,f}$.[30] Then, we collect all the markets that firms choose not to serve with direct flights and that have a similar change in the congestion costs. These markets give us a vector of lower bounds. We take the 5th percentile of these lower bounds and use it as a lower bound for $\eta_{ab,f}$. Finally, we set $\eta_{ab,f}$ as the mid-point between the lower and upper bounds. A symmetric procedure is implemented when imputing the fixed cost shocks for the markets that are not served by airline f. However, instead of the 5th percentile, we take the 95th percentile. When simulating the merger, the merged entity takes on the mean value of the fixed cost shocks of the merging firms.

To perform counterfactuals, we also need to measure the demand and marginal cost shocks of the products of the merged firm. We draw these from the joint distribution of the merged entity.

[30]Consider market $\{a, b\}$ and airline f. Suppose that $G_{ab,f} = 1$ in the observed network. Let

$$\Delta\Pi_{(-ab),f} - \Delta\overline{\mathrm{FC}}^{\top}_{(-ab),f}\gamma - \eta_{ab,f},$$

be the deviation profits of firm f as discussed in Section 2.4.2, for any $\gamma \in \Gamma_I$. By revealed preference, it must be that

$$\Delta\Pi_{(-ab),f} - \Delta\overline{\mathrm{FC}}^{\top}_{(-ab),f}\gamma - \eta_{ab,f} \geq 0,$$

or, equivalently,

$$\eta_{f,ab} \leq \Delta\Pi_{(-ab),f} - \Delta\overline{\mathrm{FC}}^{\top}_{(-ab),f}\gamma.$$

Results

We show the impact of the merger and bankruptcy scenarios in Table 2.6. The first row gives the total consumer surplus, the second row gives the average consumer surplus across hub markets. The second column, under fixed networks, provides the interval between the base-case, the best-case, and the update scenarios and the median value. The other columns, under reoptimised networks, provide the interval across draws of parameter values from the estimated identified set and the median value. In the following, we base our discussion on the median values.

Table 2.6: Consumer surplus across different scenarios

	Before	**Merger**			**Bankruptcy**	
		Networks fixed	Networks vary, no remedies	Networks varies, with remedies	Networks fixed	Networks vary
Total	2807.06	+0.08	-2.94	+2.15	-12.1	-5.5
		[-0.47, +3.4]	[-8.18, +2.28]	[-3.18, +6.61]		[-9.55, +0.97]
Mean	4.09	+0.08	-4.1	+0.88	-11.84	-5.78
		[-0.47, +3.4]	[-9.11, +0.96]	[-4.57, +5.23]		[-9.55, -0.05]

Note:
Consumer surplus is computed using the log-sum formula and it is in USD 1 million up to constant of integration. Mean consumer surplus is total consumer surplus divided by the number of markets out of hubs. Percentage differences with respect to Before are reported.

When comparing the "Networks fixed" to the "Networks vary" scenarios, we see that assuming no changes in networks after a merger or bankruptcy leads to misleading conclusions. In the merger case, the "Networks fixed" scenario predicts little changes in consumer surplus. However, when the firms are allowed to reoptimise their networks, we see that the merger absent any remedies would have led to a drop in total consumer surplus of around 2.94%. When the remedies are further taken into account, we register an increase in consumer surplus of around 2.15%, which highlights that the remedies were warranted. The difference between the "Networks fixed" and the "Networks vary" scenario is even stronger in the bankruptcy case. Here, not allowing for network re-alignment leads to a prediction of a drop in consumer surplus o around 12%, whereas letting networks re-adjust reduces the loss in consumer surplus to around 5.5%. The main reason for the smaller drop in consumer surplus is that firms enter market following the disappearance of American, partially canceling out the negative effects of the bankruptcy.

Comparing the effects of the merger and the bankruptcy across the scenarios in which we make use of our full model, we can see that the bankruptcy would have hurt consumer more than the merger. Two reasons can explain this fact. First, the firm disappearing operates a hub-and-spoke network, so consumers will be hurt a lot in hub airports and other airlines cannot compensate for the loss of access to this network. Second, a merger allows competition authorities to shape market structure and outcomes by imposing remedies. The ability to do so presents a non-negligible advantage of allowing a merger under conditions, compared to the bankruptcy of a large hub-and-spoke carrier. We see that our counterfactual suggests that the merger is less harmful than the bankruptcy absent remedies and even slightly beneficial to consumers with remedies taken into account.

Table 2.7: Changes in direct flights offered at hub airports of merging firm due to the merger

	Pre-merger			Post-merger					
				Networks vary, no remedies			**Networks vary, remedy**		
	AA/US	Others	Presence	AA/US	Others	Presence	AA/US	Others	Presence
AA hubs									
DFW	68	55	1.6	70	59	1.56	70	59	1.56
				[59, 70]	[53, 89]	[1.48, 1.84]	[60, 70]	[53, 89]	[1.48, 1.83]
LAX	28	90	1.51	27	128	1.9	30	128	1.95
				[24, 29]	[104, 132]	[1.64, 1.98]	[29, 31]	[114, 133]	[1.78, 2]
ORD	59	129	2.35	58	98	1.9	62	144	2.55
				[56, 61]	[90, 161]	[1.82, 2.67]	[61, 66]	[96, 161]	[1.93, 2.71]
MIA	40	51	1.17	18	52	0.84	40	51	1.11
				[13, 22]	[50, 57]	[0.79, 0.95]	[40, 40]	[49, 58]	[1.09, 1.2]
JFK	41	113	2	29	118	1.76	43	118	1.96
				[11, 32]	[113, 159]	[1.57, 2.29]	[43, 44]	[112, 158]	[1.89, 2.46]
US hubs									
CLT	61	41	1.29	63	35	1.2	63	33	1.17
				[61, 64]	[30, 44]	[1.11, 1.32]	[63, 64]	[24, 44]	[1.06, 1.32]
PHX	41	74	1.49	41	64	1.28	41	62	1.26
				[40, 41]	[38, 74]	[0.96, 1.39]	[41, 41]	[38, 73]	[0.96, 1.39]
DCA	40	130	2.16	49	153	2.39	39	141	2.24
				[39, 60]	[133, 161]	[2.2, 2.55]	[10, 51]	[132, 160]	[1.82, 2.48]
PHL	52	53	1.33	50	53	1.27	55	54	1.33
				[43, 52]	[49, 59]	[1.15, 1.36]	[54, 56]	[50, 64]	[1.28, 1.46]
Total									
Total	430	736	1.66	409	797	1.63	444	783	1.67
				[350, 422]	[690, 880]	[1.48, 1.77]	[403, 460]	[696, 906]	[1.54, 1.82]

Note:
Median outcomes reported, with minimum aand maximum outcome in brackets.

Table 2.8: Changes in direct flights offered at hub airports of American due to the bankruptcy

	Before			After bankruptcy	
	AA	Others	Presence	Others	Presence
AA hubs					
DFW	68	63	1.6	75	0.91
				[75, 110]	[0.91, 1.34]
LAX	26	96	1.51	140	1.73
				[134, 155]	[1.65, 1.91]
ORD	59	134	2.35	168	2.05
				[159, 172]	[1.94, 2.1]
MIA	39	57	1.17	59	0.72
				[56, 72]	[0.68, 0.88]
JFK	29	135	2	149	1.82
				[142, 190]	[1.73, 2.32]
Total					
Total	246	973	1.66	1082	1.47
				[1012, 1224]	[1.38, 1.66]

Tables 2.7 and 2.8 show the changes in the number of direct flights offered before and after the merger and bankruptcy, respectively. Looking at the last row in Table 2.7, we see that the merged entity reduces the operations at the hubs in reaction to the merger, whereas other firms expand at the hubs. The variable Presence in columns 3, 6, and 9 reports the average number of main carriers (American, Delta, United, US, Southwest) present across all possible markets out of a given hub. The value of the variable Presence drops slightly from 1.66 to 1.63 after the merger, suggesting that the merger leads to slightly less competition in hub markets. When looking at what happens in specific cities, we can see that in the "No remedies" scenario, the merged entity reduces operations in Miami (MIA) and New York (JFK) substantially, without substantially entry of the competing airlines. In the "With remedies" scenario, the merged entity serves a larger network compared to both before the merger and to the "No remedies" scenario. The expansion of other carriers is less substantial than in the "No remedies" scenario. However, the value of the variable "Presence" is now 1.67, essentially as before the merger. Overall, Table 2.7 suggests that the remedies were successful in preventing a large reduction in the post-merger network of the merged entity.

In contrast, Table 2.8 shows that post-bankruptcy, the level of operations at American's hubs decreases substantially. Even though the remaining firms increase the number of direct flights by more than 100, they are not able to compensate for the disappearance of American (with the exception of Los Angeles). The value of the variable "Presence" confirms this finding, as it drops from 1.66 to 1.47.

Table 2.9: Change in consumer surplus at hub airports of merging firms

	Pre-merger	Post-merger		
		Networks fixed	Networks vary, no remedies	Networks vary, remedy
AA hubs				
DFW	341.22	-1.48	+6.28	+8.28
		[-2.94, +7.04]	[+1.12, +15.15]	[+3.28, +16.65]
LAX	520.29	+0.01	+6.06	+9.72
		[-0.32, +2.44]	[-2.44, +7.76]	[+3.17, +11.33]
ORD	485.16	+0.46	-16.17	+2.84
		[-0.29, +4.07]	[-18.34, +4.2]	[-14.52, +6.71]
MIA	314.55	-0.34	-29.68	-19.71
		[-0.51, +4.56]	[-31.76, -26.24]	[-20.68, -17.1]
JFK	631.27	-0.3	-21.65	-14.39
		[-0.43, +2.19]	[-26.6, -11.63]	[-15.55, -4.7]
US hubs				
CLT	134.27	-1.52	+10.29	+7.81
		[-2.56, +3.27]	[+3.3, +14.43]	[+1.79, +12.44]
PHX	237.55	-0.64	-20.99	-19.3
		[-2.48, +3.66]	[-33.98, -19.12]	[-32.53, -17]
DCA	428.19	-0.29	+12.68	+12.72
		[-0.62, +2.26]	[+7, +17.07]	[+1.43, +18.49]
PHL	213.55	-0.91	+1.31	+8.48
		[-1.01, +2.98]	[-6.46, +7.43]	[+2.14, +13.58]

Note:
Consumer surplus is computed using the log-sum formula and it is in USD 1 million up to constant of integration. Mean consumer surplus is total consumer surplus divided by the number of markets out of hubs. Percentage differences with respect to Before are reported.

Tables 2.9 and 2.10 show changes in consumer surplus at the hub level. We see that the impact of the merger is quite heterogeneous across hubs, with consumers in some cities benefiting a lot and others suffering a lot. Not surprisingly, Miami and New York are among the hardest hit, mainly due to reduced operations of the merged entity. Similarly, consumer surplus drops substantially in Chicago (ORD) Phoenix (PHX). All three cities were targeted by the remedies, suggesting the concern of the state Attorney Generals was warranted. When comparing the "No remedies" scenario to the "With remedies" scenario, we see that the drop in consumer surplus in Miami and New York becomes less pronounced and even turns into a gain in Chicago, suggesting that the remedies were helpful in at least mitigating consumer harm.

Table 2.10: Change in consumer surplus at hub airports of American due to the bankruptcy

	Before	After bankruptcy	
		Networks fixed	Networks vary
AA hubs			
DFW	341.22	-34.37	-16.4
			[-17.37, +0.13]
LAX	520.29	-10.86	+7.1
			[+3.85, +11.8]
ORD	485.16	-16.85	-11.37
			[-16.71, -10.05]
MIA	314.55	-15.33	-30.2
			[-32.61, -25.39]
JFK	631.27	-9.13	-18.36
			[-21.04, -7.68]

Note:
Consumer surplus is computed using the log-sum formula and it is in USD 1 million up to constant of integration. Mean consumer surplus is total consumer surplus divided by the number of markets out of hubs. Percentage differences with respect to Before are reported.

In contrast, as Table 2.10 suggests, the picture is bleaker in the case of American's bankruptcy. With the exception of Los Angeles - where other firms enter a lot- consumer surplus drops substantially. Again, Miami and New York suffer the most. It is also noteworthy that Dallas/Fort Worth (DFW), American's biggest hub, now sees consumer surplus drop by around 16%. Comparing the impact on consumer surplus on hubs between the merger and the bankruptcy of American underlines the negative effects of removing a large hub-and-spoke airline from the market: consumers suffer for no longer having access to a large network and hub amenities and other airlines struggle to fill the void left behind by the bankruptcy.

Finally, Table 2.11 shows the merger's and bankruptcy's impact on prices, markups, and marginal costs. We observe a drop in prices in response to both the merger and the bankruptcy (with the exception of one-stop flights offered by other airlines). At the same time, marginal cost also falls quite a bit. For the merging entity, the drop in marginal cost is due to a larger network post-merger. Even though their post-merger network is smaller than their combined pre-merger network, the post-merger network is still larger than the individual per-merger networks, allowing for marginal cost savings. Similarly, the other carriers increase their operations substantially both in response to a merger and a bankruptcy, leading to marginal cost savings from a larger network and to lower prices. Interestingly, the remedies reduce the increase in

Table 2.11: Changes in prices, marginal cost, and markups

	Before	After		
		Merger		**Bankruptcy**
		No remedies	With remedies	
AA: Direct				
Price	406.24	-6.67	-6.56	
		[-6.96, -6.15]	[-7.08, -5.93]	
Marginal Cost	276.7	-11.54	-10.76	
		[-11.82, -11.29]	[-11.19, -9.5]	
Markup	129.54	+3.84	+2.53	
		[+3, +5.31]	[+0.52, +3.71]	
Others: Direct				
Price	413.19	-1.63	-4.01	-2.85
		[-4.18, +0.24]	[-4.98, -0.77]	[-3.9, -1.88]
Marginal Cost	291.6	-4.42	-6.38	-5.68
		[-7.61, -1.14]	[-7.95, -1.77]	[-7.28, -4.37]
Markup	121.59	+3.85	+1.74	+3.73
		[+3.23, +5.06]	[+1.37, +2.69]	[+3.42, +4.1]
AA: One-stop				
Price	466.39	-6.37	-7.06	
		[-7.2, -5.4]	[-7.58, -6.68]	
Marginal Cost	351.28	-10.94	-11.4	
		[-12.15, -9.57]	[-12.05, -10.42]	
Markup	115.11	+8.1	+6.51	
		[+5.51, +9.36]	[+3.28, +7.4]	
Others: One-stop				
Price	416.12	+2.41	+1.75	+1.47
		[+1.32, +3.17]	[+0.97, +2.6]	[+1.13, +1.82]
Marginal Cost	301.18	+2.71	+2.53	+1.24
		[+1.24, +3.81]	[+1.35, +3.65]	[+0.71, +1.66]
Markup	114.94	+1.1	-0.26	+2.18
		[+0.7, +2.75]	[-0.74, +0.82]	[+1.91, +2.62]

markups for all firms and products. The reason may be increased competition, especially since American faces pressure to raise markups as the remedies lead to a larger network with higher fixed costs.

Discussion

Overall, our counterfactual exercise suggests that the merger had a small positive impact on consumer surplus. However, the overall effect hides substantial heterogeneity across cities: consumers in some hub markets saw a large decrease in consumer surplus that was mitigated, but not reversed by the remedies imposed. Further, our results suggest that letting American and US Airways merge was better than the bankruptcy and disappearance of American. Even though the remaining firms enter substantially in response to American disappearing, they cannot completely fill out the void left behind. We also find that the remedies that the merged entity agreed to and put a floor on levels of operation at the majority of hubs helped in mitigating harm to consumers and pushed the overall change in consumer surplus from slightly negative to slightly positive. The positive effect of these remedies shows an advantage of mergers over the disappearance of firms in especially hub-and-spoke networks: competition authorities can shape post-merger outcomes by imposing remedies.

2.6 Conclusions

We consider a two-stage model of airline competition where airlines design their route networks in the first stage and compete in prices in the second stage. We show identification of the second-stage parameters by following the standard approach for supply and demand models with differentiated products. We show (set) identification of the first-stage parameters by adopting a revealed preference perspective and exploiting inequalities derived from equilibrium implications. We estimate our model using data on the US airline industry from the second quarter of 2011. In the first stage, we find that fixed costs increase in the number of destinations reachable from hub airports. On the supply side of the second stage, we find that marginal costs decrease in the number of flights (direct or one-stop) offered out of the endpoints. On the demand side of the second stage, we find that consumer utility increases in the number of direct connections that can be reached from the endpoints. We then use the results to evaluate the merger between American Airlines and US Airways which did occur at a later date. We find that remedies imposed on the merging parties turned a slight decrease in consumer surplus into a slight increase. At the most negatively affected hubs, the remedies helped to contain consumer harm. We also compare the merger to a hypothetical bankruptcy and disappearance of American Airlines. We find that the bankruptcy leads to more rival firm entry than the merger. However, the loss of access to a large network leads to substantial loss in consumer surplus at American's hubs. Other firms are not able to fill this void completely. Overall, consumer surplus is projected to decrease by more than in the merger case.

Our work leaves several avenues for future research. In our model, we abstract from capacity and frequency choices which are an important point of concern both to consumers as well as antitrust authorities. Extending our framework to include these kind of choices is possible, albeit at the cost of increasing the computational burden. Similarly, endogenising hub decisions would make it possible to analyse deeper changes in network structure, such as the choice between a hub-and-spoke and a point-to-point network. Assuming exogenous hubs prevents firms from creating new hubs, which may be especially interesting to consider in the case where we let American Airlines disappear. However, we believe that the assumption of exogenous hubs is reasonable in our setting for two reasons: First, the other airlines already have large hub-and-spoke networks, making it very costly to create additional hubs. Second, the trend in the US Airline industry has been to reduce the number of hubs, rather than increase them (see also Berry and Jia, 2010). Finally, we do not consider all remedies that were imposed on the new merged entity. For instance, the US Department of Justice ordered the merged entity to give up slots and gates at several airports in order to facilitate the entry of new airlines. Adding capacity and frequency choices to our model would allow for a detailed evaluation of these remedies.

We are currently working on inference for the first-stage parameters and on further robustness checks for the counterfactual part.

References

Jason Abaluck and Abi Adams. What do consumers consider before they choose? Identification from asymmetric demand responses. *Quarterly Journal of Economics (forthcoming)*, 2020.

Victor Aguirregabiria and Chun-Yu Ho. A dynamic oligopoly game of the US airline industry: Estimation and policy experiments. *Journal of Econometrics*, 168(1):156–173, 2012.

Donald WK Andrews and Soonwoo Kwon. Inference in moment inequality models that is robust to spurious precision under model misspecification. Working Paper, 2019.

Isaiah Andrews, Jonathan Roth, and Ariel Pakes. Inference for linear conditional moment inequalities. NBER Working Papers 26374, 2019.

Levon Barseghyan, Maura Coughlin, Francesca Molinari, and Joshua C Teitelbaum. Heterogeneous choice sets and preferences. *Econometrica (forthcoming)*, 2020.

C Lanier Benkard, Aaron Bodoh-Creed, and John Lazarev. Simulating the Dynamic Effects of Horizontal Mergers: U.S. Airlines. Working Paper, 2020.

Arie Beresteanu and Francesca Molinari. Asymptotic properties for a class of partially identified models. *Econometrica*, 76(4):763–814, 2008.

Steven T Berry. Airport presence as product differentiation. *American Economic Review*, 80 (2):394–399, 1990.

Steven T Berry. Estimation of a model of entry in the airline industry. *Econometrica*, 60(4): 889–917, 1992.

Steven T Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262, 1994.

Steven T Berry and Philip A Haile. Identification in differentiated products markets using market level data. *Econometrica*, 82(5):1749–1797, 2014.

Steven T Berry and Panle Jia. Tracing the woes: An empirical analysis of the airline industry. *American Economic Journal: Microeconomics*, 2(3):1–43, 2010.

Steven T Berry and Peter Reiss. Empirical models of entry and market structure. *Handbook of Industrial Organization*, 3:1845–1886, 2007.

Steven T Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica*, 63(4):841–890, 1995.

Steven T Berry, Michael Carnall, and Pablo T Spiller. Airline hubs: costs, markups and the implications of customer heterogeneity. NBER Working Papers 5561, 1996.

Steven T Berry, Oliver B Linton, and Ariel Pakes. Limit theorems for estimating the parameters of differentiated product demand systems. *The Review of Economic Studies*, 71(3):613–654, 2004.

Steven T Berry, Martin Gaynor, and Fiona Scott Morton. Do increasing markups matter? Lessons from empirical industrial organization. *Journal of Economic Perspectives*, 33(3): 44–78, 2019.

Christian Bontemps, Thierry Magnac, and Eric Maurin. Set identified linear models. *Econometrica*, 80(3):1129–1155, 2012.

Severin Borenstein. Hubs and high fares: dominance and market power in the U.S. airline industry. *The RAND Journal of Economics*, 20(3):344–365, 1989.

Severin Borenstein. The evolution of U.S. airline competition. *Yale Journal on Regulation*, 4 (2):45–73, 1992.

Timothy F Bresnahan. Competition and collusion in the American automobile industry: The 1955 price war. *The Journal of Industrial Economics*, 35(4):457–482, 1987.

Timothy F Bresnahan and Peter C Reiss. Entry in monopoly market. *The Review of Economic Studies*, 57(4):531–553, 1990.

Timothy F Bresnahan and Peter C Reiss. Empirical models of discrete games. *Journal of Econometrics*, 48(1):57–81, 1991.

Jan K Brueckner and Pablo T Spiller. Economies of traffic density in the deregulated airline industry. *The Journal of Law & Economics*, 37(2):379–415, 1994.

Jan K Brueckner, Nichola J Dyer, and Pablo T Spiller. Fare determination in airline hub-and-spoke networks. *The RAND Journal of Economics*, 23(3):309–333, 1992.

Richard V Butler and John H Houston. How contestable are airline markets. *Atlantic Economic Journal*, 17:27–35, 1989.

Kenneth Button, Peter Forsyth, and Peter Nijkamp. *Air Transport*. Edward Elgar, 2000.

Douglas W Caves, Laurits R Christensen, and Michael W Tretheway. Economies of density versus economies of scale: why trunk and local service airline costs differ. *The RAND Journal of Economics*, 15(4):471–489, 1984.

Arun Chandrasekhar. Econometrics of network formation. In *The Oxford Handbook of the Economics of Networks*, pages 303–357. Oxford University Press Oxford, 2016.

Arun G Chandrasekhar, Victor Chernozhukov, Francesca Molinari, and Paul Schrimpf. Best linear approximations to set identified functions: with an application to the gender wage gap. NBER Working Papers 25593, 2019.

Federico Ciliberto and Elie Tamer. Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828, 2009.

Federico Ciliberto and Jonathan W Williams. Does multimarket contact facilitate tacit collusion? Inference on conduct parameters in the airline industry. *The RAND Journal of Economics*, 45(4):764–791, 2014.

Federico Ciliberto, Charles Murry, and Elie Tamer. Market structure and competition in airline markets. *Journal of Political Economy (forthcoming)*, 2021.

Somnath Das. Effect of merger on market price and product quality: American and US Airways. *Review of Industrial Organisation*, 55:339–374, 2019.

Aureo de Paula. Econometrics of network models. In *Advances in Economics and Econometrics: Theory and Applications, Eleventh World Congress*, pages 268–323. Cambridge University Press Cambridge, 2017.

Aureo de Paula. Econometric models of network formation. *Annual Review of Economics*, 12: 775–799, 2020.

Alon Eizenberg. Upstream Innovation and Product Variety in the U.S. Home PC Market. *Review of Economic Studies*, 81(3):1003–1045, 2014.

Joseé-Antonio Espín-Sánchez, Álvaro Parra, and Yuzhou Wang. Equilibrium uniqueness in entry games with private information. Working Paper, 2021.

William N Evans and Ioannis N Kessides. Localized market power in the US airline industry. *The Review of Economics and Statistics*, 75(1):66–75, 1993.

Xiaowen Fu, Huan Jin, Shaoxuan Liu, Tae H Oum, and Jia Yan. Exploring network effects of point-to-point networks: An investigation of the spatial patterns of southwest airlines' network. *Transport Policy*, 76:36–45, 2019.

Bulat Gafarov. Inference in high-dimensional set-identified affine models. arXiv Working Papers 1904.00111, 2019.

Amit Gandhi, Zhentong Lu, and Xiaoxia Shi. Estimating demand for differentiated products with zeroes in market share data. Working Paper, 2019.

Austan Goolsbee and Chad Syverson. How do incumbents respond to the threat of entry? Evidence from the major airlines. *Quarterly Journal of Economics*, 123(4):1611–1633, 2008.

Bryan Graham and Aureo de Paula. *The Econometric Analysis of Network Data*. Academic Press, 2020.

Bryan S Graham. Methods of identification in social networks. *Annual Review of Economics*, 7(1):465–485, 2015.

Cristina Gualdani. An econometric model of network formation with an application to board interlocks between firms. *Journal of Econometrics (forthcoming)*, 2021.

Tim Hellmann. On the existence and uniqueness of pairwise stable networks. *International Journal of Game Theory*, 42:211–237, 2013.

Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex Analysis and Minimization Algorithms*. Springer-Verlag, 1996.

Thomas J Holmes. The diffusion of wal-mart and economies of density. *Econometrica*, 79(1): 253–302, 2011.

Jean-François Houde, Peter Newberry, and Katja Seim. Economies of Density in E-Commerce: A Study of Amazon's Fulfillment Center Network. NBER Working Papers 23361, 2017.

Matthew O Jackson and Alison Watts. The existence of pairwise stable networks. *Seul Journal of Economics*, 14(3):299–321, 2001.

Matthew O Jackson and Alison Watts. The evolution of social and economic networks. *Journal of Economic Theory*, 106(2):265–295, 2002.

Matthew O Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71(1):44–74, 1996.

Hiroaki Kaido and Andres Santos. Asymptotically efficient estimation of models defined by convex moment inequalities. *Econometrica*, 82(1):387–413, 2014.

Adib Kanafani and Atef A Ghobrial. Airline hubbing - Some implications for airport economics. *Transportation Research Part A: General*, 19(1):15–27, 1985.

Joseph Kuehn. Spillovers from entry: the impact of bank branch network expansion. *The RAND Journal of Economics*, 49(4):964–994, 2018.

Robin S Lee and Ariel Pakes. Multiple equilibria and selection by learning in an applied setting. *Economics Letters*, 104(1):13–16, 2009.

Micheal P Leung. Dependence-robust inference using resampled statistics. Working Paper, 2020.

Micheal P Leung. Network cluster-robust inference. Working Paper, 2021.

Michael E Levine. Airline competition in deregulated markets: theory, firm strategy, and public policy. *Journal of Economic Perspective*, 6:393–494, 1987.

Sophia Ying Li, Joe Mazur, Yongjoon Park, James W Roberts, Andrew Sweeting, and Jun Zhang. Repositioning and Market Power After Airline Mergers. Working Paper, 2019.

Israel Mark, Bryan Keating, Daniel L Rubinfeld, and Bobby Willig. Airline network effects and consumer welfare. *Review of Network Economics*, 12(3):287–322, 2013.

Antonio Mele. A structural model of dense network formation. *Econometrica*, 85(3):825–850, 2017.

Yuhei Miyauchi. Structural estimation of pairwise stable networks with nonnegative externality. *Journal of Econometrics*, 195(2):224–235, 2016.

Dov Monderer and Lloyd S Shapley. Potential games. *Games and Economic Behavior*, 14(1): 124–143, 1996.

Steven A Morrison and Clifford Winston. *The economic Effects of Airline Deregulation*. Washington, DC: Brookings Institution, 1986.

Steven A Morrison and Clifford Winston. Enhancing the performance of the deregulated air transport system. *Brookings Papers on Economic Activity. Microeconomics*, 1989:61–123, 1989.

Giovanni Nero. A note on the competitive advantage of large hub-and-spoke networks. *Transportation Research Part E: Logistics and Transportation Review*, 35(4):225–239, 1999.

Volker Nocke and Nicolas Schutz. Multiproduct-firm oligopoly: An aggregative games approach. *Econometrica*, 86(2):523–557, 2018.

Tae Hoon Oum, Anming Zhang, and Yimin Zhang. Airline network rivalry. *Canadian Journal of Economics*, 28(4a):836–857, 1995.

Ariel Pakes. Alternative models for moment inequalities. *Econometrica*, 78(6):1783–1822, 2010.

Ariel Pakes, Jack Porter, Kate Ho, and Joy Ishii. Moment inequalities and their application. *Econometrica*, 83(1):315–334, 2015.

Yongjoon Park. Structural remedies in network industries: An assessment of slot divestitures in the American Airlines/US Airways merger. Working Paper, 2020.

Craig Peters. Evaluating the performance of merger simulation: Evidence from the U.S. airline industry. *The Journal of Law & Economics*, 49(2):627–649, 2006.

Peter C Reiss and Pablo T Spiller. Competition and entry in small airline markets. *The Journal of Law & Economics*, 32(2):S179–S202, 1989.

Aisling Reynolds-Feighan. Traffic distribution in low-cost and full-service carrier networks in the US air transportation market. *Journal of Air Transport Management*, 7(5):265–275, 2001.

Joseph Anthony Rossetti. *Product Variety in the US Yogurt Industry*. PhD thesis, The Ohio State University, 2018.

Alexander Shapiro, Darinka Dentcheva, and Andrzej Ruszczyński. *Lectures on Stochastic Programming: Modelling and Theory*. SIAM, 2014.

Shuyang Sheng. A structural econometric analysis of network formation games through subnetworks. *Econometrica*, 88(5):1829–1858, 2020.

Jörg Stoye. More on confidence intervals for partially identified parameters. *Econometrica*, 77 (4):1299–1315, 2009.

Donald M Topkis. Equilibrium points in nonzero-sum n-person submodular games. *SIAM Journal of Control and Optimization*, 17(6):773–787, 1979.

Thomas G Wollmann. Trucks without bailouts: Equilibrium product characteristics for commercial vehicles. *American Economic Review*, 108(6):1364–1406, 2018.

Zhe Yuan. Network competition in the airline industry: A framework for empirical policy analysis. Working Paper, 2020.

Appendices

Appendix F

Existence of Nash equilibrium networks

As discussed in Section 2.3.3, our methodology does not require the existence of Nash equilibrium networks for every possible parameterization and realization of the variables. Formally proving the existence of Nash equilibrium networks is a difficult and open theoretical question, due to the presence of spillover effects in entry decisions across markets on the demand, marginal cost, and fixed cost sides.

Berry (1992) establishes equilibrium existence in one of the first empirical models of entry that accounts for the strategic interactions among the airlines in the second-stage pricing game. His proof relies on the assumption that the entry decisions are independent across markets and, hence, it is not applicable to our case. Another approach that has been used in the literature to show existence of Nash equilibrium networks consists of representing the model as a potential game (Monderer and Shapley, 1996). This seems to be a feasible exercise only when the payoff function is additive separable in the linking decisions and linear in the spillover effects (for example, Mele, 2017), which is not the case here. Alternatively, it is possible to show the existence of Nash equilibrium networks by assuming that the game is supermodular (for example, Miyauchi, 2016; Sheng, 2020), in order to rely on the fixed-point theorem for isotone mappings (Topkis, 1979). However, supermodularity does not hold in our setting due to the second-stage competition among the airlines. Finally, one could attempt to decompose the original game into "local" games such that the original game is in equilibrium if and only if each local game is in equilibrium (Gualdani, 2021). In turn, the existence of an equilibrium in each local game - which is typically easier to be established - is sufficient for the existence of an equilibrium in the original game. However, the classes of spillover effects considered in our model do not allow us to implement such a decomposition.

It should also be noticed that the revealed-preference inequalities, which we use to bound the first-stage parameters, resemble the notion of pairwise stability used in network theory, where no players have profitable deviations by adding or removing a link (Jackson and Wolinsky, 1996). Therefore, we have explored the possibility of establishing an equilibrium in entry decisions weaker than Nash equilibrium, along the lines of pairwise stability. In particular, according to Jackson and Watts (2002), for any payoff function there is either a pairwise stable network or a closed cycle.[1] A typical way used in the literature to exclude the presence of closed cycles consists of showing that the model can be represented as a potential game, as discussed in Jackson and Watts (2001) and Hellmann (2013). As earlier, however, this requires the payoff function to be additive separable in the linking decisions and linear in the spillover effects (for example, Sheng, 2020), which is not our case.[2]

[1]A closed cycle represents a situation in which individuals never reach a stable state and constantly switch between forming and severing links.

[2]One may wonder whether allowing for *private* fixed cost shocks could simplify the existence proof. Espín-Sánchez, Parra, and Wang (2021) prove equilibrium existence in a class of entry model where the firms have some private information at the entry stage. However, the class of entry models they consider do not allow for multi-product firms and for spillover effects in entry decisions across markets. Further, in our setting, we view more reasonable to assume that the fixed cost shocks are common knowledge among the airlines, as discussed in Section 2.3.2.

Appendix G

Inference

G.1 Inference on the second-stage parameters

We conduct inference on θ_0 via GMM and under the assumption that the number of markets goes to infinity. Formally, we consider the moment conditions of Section 2.4.1 and use their sample analogues to construct a GMM objective function which should be minimised with respect to $\theta \in \Theta$:

$$Q(\theta) = M(\theta)' A M(\theta), \tag{G.1.1}$$

where

$$M(\theta) \equiv \begin{pmatrix} \frac{1}{|\mathcal{J}|}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}_t}[\tau_{j,t}(X_t^{\oplus}, W_t^{\oplus}, M, s_t^{\oplus}, P_t^{\oplus}, G;\theta_0) \times z_{j,t,1}(X_t^{\oplus}, W_t^{\oplus})] \\ \frac{1}{|\mathcal{J}|}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}_t}[\tau_{j,t}(X_t^{\oplus}, W_t^{\oplus}, M, s_t^{\oplus}, P_t^{\oplus}, G;\theta_0) \times z_{j,t,2}(X_t^{\oplus}, W_t^{\oplus})] \\ \vdots \\ \frac{1}{|\mathcal{J}|}\sum_{t\in\mathcal{T}}\sum_{j\in\mathcal{J}_t}[\tau_{j,t}(X_t^{\oplus}, W_t^{\oplus}, M, s_t^{\oplus}, P_t^{\oplus}, G;\theta_0) \times z_{j,t,\mathrm{L}}(X_t^{\oplus}, W_t^{\oplus})] \end{pmatrix},$$

$\mathcal{J} \equiv \cup_{t\in\mathcal{T}}\mathcal{J}_t$ is the set of all offered products, and A is an appropriate $2\mathrm{L} \times 2\mathrm{L}$ weighting matrix. In particular, A is computed via the usual two-step procedure: first, we estimate the parameters using the optimal weighting matrix under conditional homoskedasticity; second, we use the obtained estimates to construct the optimal weighting matrix under conditional heteroskedasticity and re-estimate the parameters.

Note that we estimate the demand and supply sides jointly. We could also estimate the demand and supply sides separately, by following a two-step procedure: first, the demand parameters are estimated; then, these estimates are used to compute the markups; lastly, the resulting marginal costs are regressed on the observed marginal cost shifters to obtain the supply parameters. We have decided to estimate the demand and supply sides jointly because it allows us to take into account the potential correlation between the demand and supply moments and, hence, obtain more precise estimates, as discussed in Berry et al. (1995). Further, given that we have a computationally "light" demand specification, the additional cost of estimating the demand and supply sides jointly is negligible.

Finally, we can account for the non i.i.d.ness of observations across markets by using HAC or cluster-robust standard errors (Leung, 2021).

G.2 Inference on the first-stage parameters

Following Section 2.4.2, the vector of first-stage parameters, γ_0, is set identified by $\mathrm{N}\times(\mathrm{R}_{+}+\mathrm{R}_{-})$ moment inequalities, where R_{+} is the number of instruments available for the class of deviations "$(+ab)$" for each firm, R_{-} is the number of instruments available for the class of deviations "$(-ab)$" for each firm, and N is the number of firms. These moment inequalities are

$$
\begin{aligned}
&\mathbb{E}[Z^{r_+}_{(+ab),f} \times (\Delta\Pi_{(+ab),f} - \Delta\overline{\mathrm{FC}}^{\top}_{(+ab),f}\ \gamma_0)|G_{ab,f}=0] \geq 0,\\
&\mathbb{E}[Z^{r_-}_{(-ab),f} \times (\Delta\Pi_{(-ab),f} - \Delta\overline{\mathrm{FC}}^{\top}_{(-ab),f}\ \gamma_0)|G_{ab,f}=1] \geq 0,\\
&r_+ = 1,\dots,\mathrm{R}_+, \quad r_- = 1,\dots,\mathrm{R}_-, \quad f = 1,\dots,\mathrm{N},
\end{aligned}
$$

where $Z^{r_+}_{(+ab),f}, Z^{r_-}_{(-ab),f}$ are the instruments, $\Delta\Pi_{(+ab),f}, \Delta\Pi_{(-ab),f}$ are the differences in the expected second-stage profits, and $\Delta\overline{\mathrm{FC}}^{\top}_{(+ab),f}\ \gamma_0, \Delta\overline{\mathrm{FC}}^{\top}_{(-ab),f}\ \gamma_0$ are the differences in the systematic fixed costs. Further, it is useful to rewrite the above moment inequalities as unconditional moment inequalities,

$$
\begin{aligned}
&\mathbb{E}[(1-G_{ab,f}) \times Z^{r_+}_{(+ab),f} \times (\Delta\Pi_{(+ab),f} - \Delta\overline{\mathrm{FC}}^{\top}_{(+ab),f}\ \gamma_0)] \geq 0,\\
&\mathbb{E}[G_{ab,f} \times Z^{r_-}_{(-ab),f} \times (\Delta\Pi_{(-ab),f} - \Delta\overline{\mathrm{FC}}^{\top}_{(-ab),f}\ \gamma_0)] \geq 0,\\
&r_+ = 1,\dots,\mathrm{R}_+, \quad r_- = 1,\dots,\mathrm{R}_-, \quad f = 1,\dots,\mathrm{N}.
\end{aligned}
\tag{G.2.1}
$$

The moment inequalities in (G.2.1) are linear in γ_0. Therefore, the identified set for γ_0, Γ_I, is a convex polytope. We assume that Γ_I is nonempty and bounded. The non-emptiness of Γ_I means that our structural model is well-specified and the instruments are valid. Andrews and Kwon (2019) propose a test for misspecification which could be used here. The boundedness of Γ_I means that, within the classes of deviations considered, the instruments capture sufficient variations in profits relative to the support of the first-stage shocks.

Convexity has been proved to be a particularly attractive feature in the set identification literature (Beresteanu and Molinari, 2008; Bontemps, Magnac, and Maurin, 2012; Kaido and

Santos, 2014). In fact, it often reduces the computational burden of estimation because the analysts can focus on estimating the support function of the identified set. The support function of Γ_I describes the distances of the supporting hyperplanes of Γ_I in each direction from the origin (Figure G.2.1). If the chosen direction, q, has its k-th component equal to 1 (resp.,

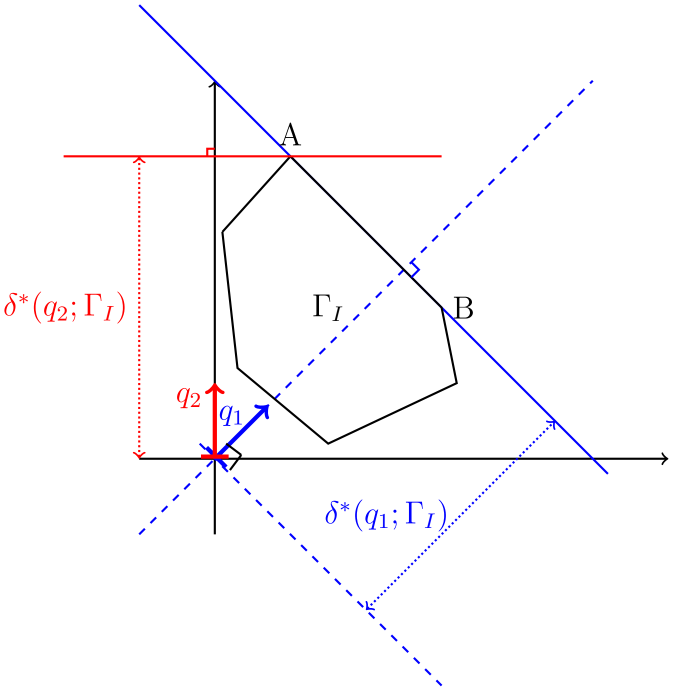


Figure G.2.1: The support function. A is a vertex. [AB] is an exposed face.

-1) and the other components equal to 0, then the support function of Γ_I in direction q is equal to the maximum (resp., minus the minimum) value of the k-th component of $\gamma \in \Gamma_I$.[1] Therefore, constructing a confidence interval for any component (or, any linear combination of components) of $\gamma \in \Gamma_I$ involves the estimation of the support function in two specific directions only.

In the next paragraphs, we elaborate on the above discussion. Our exposition is articulated in three steps. First, we argue that the support function of Γ_I can be rewritten as a linear program. Second, we derive the asymptotic distribution of the estimated support function of Γ_I. Third, we show how to use such asymptotic distribution to construct a confidence interval for any component of $\gamma \in \Gamma_I$.[2]

For easiness of exposition, in what follows we focus only on the moment inequalities of firm f

[1] In particular, one can easily contruct an outer rectangular set of Γ_I by considering these two directions for each k-th component of $\gamma \in \Gamma_I$.

[2] Andrews, Roth, and Pakes (2019) develop an inference method for a class of linear conditional moment inequalities. Our approach exploits the convexity of the identified set and allows us to easily incorporate the sampling uncertainty induced by the estimation of θ_0. Comparison with their approach is left to future research.

and on the class of deviations "$(-ab)$":

$$\mathbb{E}[G_{ab,f} \times Z^{r_-}_{(-ab),f} \times (\Delta\Pi_{(-ab),f} - \Delta\overline{\mathrm{FC}}^{\top}_{(-ab),f}\ \gamma_0)] \geq 0, \quad r_- = 1, \ldots, \mathrm{R}_-. \tag{G.2.2}$$

We streamline the notation of (G.2.2) as

$$\mathbb{E}(Z_{r,m}A_m) - \mathbb{E}(Z_{r,m}B_m^{\top})\gamma_0 \geq 0, \quad r = 1, \ldots, \mathrm{R},$$

where the subscript f is omitted; m is the market index and replaces the subscripts ab and $(-ab)$; $Z_{r,m}$ is the instrument and replaces $Z^{r_-}_{(-ab,f)}$; R is the number of instruments and replaces R_-; A_m is the difference in the expected second-stage profits multiplied by $G_{m,f}$ and replaces $G_{ab,f} \times \Delta\Pi_{(-ab),f}$; $B_m^{\top}\gamma_0$ is the difference in the systematic fixed cost multiplied by $G_{m,f}$ and replaces $G_{ab,f} \times \Delta\overline{\mathrm{FC}}^{\top}_{(-ab),f}\ \gamma_0$.

Step 1 Following Hiriart-Urruty and Lemaréchal (1996) (p.235), the support function of Γ_I in any direction q can be written as a linear program in the standard form by strong duality:

$$\begin{aligned}
\delta(q, \Gamma_I) &\equiv \sup_{\gamma \in \Gamma_I} q^{\top}\gamma, \\
&= \inf_{t \geq 0} \sum_{r=1}^{\mathrm{R}} t_r \mathbb{E}(Z_{r,m}A_m), \\
&\quad \text{s.t.} \ \sum_{r=1}^{\mathrm{R}} t_r \mathbb{E}(Z_{r,m}B_m) = q.
\end{aligned} \tag{G.2.3}$$

The Lagrangian of (G.2.3) is equal to

$$L(t, \mu, \nu) \equiv \sum_{r=1}^{\mathrm{R}} t_r \mathbb{E}(Z_{r,m}A_m) + \mu^{\top}\left(\sum_{r=1}^{\mathrm{R}} t_r \mathbb{E}(Z_{r,m}B_m) - q\right) - \nu^{\top}t, \tag{G.2.4}$$

where $\mu = (\mu_1, \ldots, \mu_{\mathrm{N}+1})$ is the vector of Lagrange multipliers for the equality constraints; $\nu = (\nu_1, \ldots, \nu_{\mathrm{R}})$ is the vector of Lagrange multipliers for the inequality constraints. We denote by $\mathcal{T}$ and $\mathcal{M}$ the sets of t and μ satisfying the KKT conditions, respectively.

Step 2 We make the simplifying assumption that θ_0 is known by the researcher. In practice, θ_0 is estimated and we explain how to account for the resulting sampling uncertainty at the end of the section. We further make the simplifying assumption that we have an i.i.d. random sample of observations

$$\{Z_{1,m}, \ldots, Z_{R,m}, A_m, B_m\}_{m=1}^{\mathrm{M}},$$

where M is the number of sampled markets, and that the Central Limit Theorem applies to all the average of the quantities of interest. In practice, our observations are not i.i.d. across markets and one can account for it, for example, by implementing the resampling approach by

Leung (2020).[3]

We introduce some notation that is useful for the next arguments. For every $r = 1, \dots, \mathrm{R}$, let X_r be the limit in distribution of $\sqrt{\mathrm{M}}(\frac{1}{\mathrm{M}}\sum_{m=1}^{\mathrm{M}} Z_{r,m}B_m - \mathbb{E}(Z_{r,m}B_m))$, i.e., a $(\mathrm{N}+1)\times 1$ random normal vector centered with variance-covariance matrix $Var(Z_{r,m}B_m)$. Fix $j = 1, \dots \mathrm{N}+1$. Select the j-th element from X_r. Repeat this operation for each $r = 1, \dots, \mathrm{R}$. Denote by X_j the resulting $\mathrm{R}\times 1$ vector. Let W_r be the limit in distribution of $\sqrt{\mathrm{M}}(\frac{1}{\mathrm{M}}\sum_{m=1}^{\mathrm{M}} Z_{r,m}A_m - \mathbb{E}(Z_{r,m}A_m))$, i.e., a random normal variable centered with variance $Var(Z_{r,m}A_m)$. As earlier, note that the random variables $\{W_r\}_{r=1}^{\mathrm{R}}$ are correlated. Let the estimated identified set be defined as

$$\widehat{\Gamma}_I \equiv \Big\{\gamma \in \Gamma : \frac{1}{\mathrm{M}}\sum_{m=1}^{\mathrm{M}} Z_{r,m}B_m^\top \gamma \leq \frac{1}{\mathrm{M}}\sum_{m=1}^{\mathrm{M}} Z_{r,m}A_m \text{ for } r = 1, \dots, \mathrm{R}\Big\}.$$

Let the estimated support function in direction q be defined as

$$\hat{\delta}(q;\Gamma_I) \equiv \delta(q;\widehat{\Gamma}_I).$$

Theorem G.2.1 provides the asymptotic distribution of $\hat{\delta}(q;\Gamma_I)$ in any direction q.

Theorem G.2.1. Assume that the moments of order $2+\tau$ of the random variables exist for some $\tau > 0$. Then:

(i) The estimated support function, $\hat{\delta}(q;\Gamma_I)$, tends to the true support function, $\delta(q;\Gamma_I)$, uniformly in q in the unit ball.

(ii) It holds that, uniformly in q,

$$\sqrt{\mathrm{M}}\left(\hat{\delta}(q;\Gamma_I) - \delta(q;\Gamma_I)\right) \xrightarrow[\mathrm{M}\to\infty]{d} \inf_{t\in\mathcal{T}} \sup_{\mu\in\mathcal{M}} \left[\sum_{r=1}^{\mathrm{R}} t_r W_r + \sum_{j=1}^{\mathrm{N}+1} \mu_j t^\top X_j\right].$$

If $\mathcal{T}$ and $\mathcal{M}$ are singleton, then the asymptotic distribution is normal. $\diamond$

Proof. (i) comes from the convergence of $\widehat{\Gamma}_I$ to Γ_I with respect to the Hausdorff distance. (ii) comes from Shapiro, Dentcheva, and Ruszczyński (2014), Theorem 5.11, p.173. $\square$

[3]Note that our methodology allows for the airlines' networks to be partially observed, provided that we fully observe the portions of the airlines' networks whose nodes are the cities at the endpoints of the sampled markets.

Step 3 Theorem G.2.1 provides the asymptotic distribution of the estimated support function in any direction. Hence, as discussed at the beginning of this section, it can be used to derive confidence regions for any component (or, any linear combination of components) of $\gamma \in \Gamma_I$. However, note that such asymptotic distribution depends on $\mathcal{T}$ and $\mathcal{M}$. If these sets are singleton, then the estimated support function is asymptotically normal, with a variance that can be estimated from the data. Unfortunately, these sets are not singleton in all directions q. For instance, consider the directions which correspond to the outer normal of an exposed face of Γ_I (e.g., q_1 in Figure G.2.1). This is a well-known problem in the set identification literature. One solution consists of smoothing Γ_I. Chandrasekhar, Chernozhukov, Molinari, and Schrimpf (2019) transform the explanatory variables into continuous ones by adding $\varepsilon N(0,1)$ to each discrete explanatory variable. Bontemps et al. (2012) and Gafarov (2019) perturb the support function by adding a penalty term.

Here, we follow Gafarov (2019)'s approach and add the penalty term $\varepsilon||\gamma||_2$ to the support function of Γ_I in any direction q:

$$\delta_\varepsilon(q,\Gamma_I) \equiv \sup_{\gamma\in\Gamma_I} q^\top\gamma - \varepsilon||\gamma||_2, \tag{G.2.5}$$

with $\epsilon > 0$ but small. Since $\varepsilon||\gamma||_2$ is strictly convex, (G.2.5) has a unique solution with respect to γ, for any direction q. In turn, by strong duality, it holds that

$$\begin{aligned} \delta_\varepsilon(q,\Gamma_I) = \inf_{t\geq 0} & \sum_{r=1}^{\mathrm{R}} t_r \mathbb{E}(Z_{r,m}A_m), \\ \text{s.t.} & \left\| \sum_{r=1}^{\mathrm{R}} t_r \mathbb{E}(Z_{r,m}B_m) - q \right\|_2 \leq \varepsilon. \end{aligned} \tag{G.2.6}$$

The Lagrangian of (G.2.6) is equal to

$$L(t,\mu,\nu) \equiv \sum_{r=1}^{\mathrm{R}} t_r \mathbb{E}(Z_{r,m}A_m) + \mu^\top \left(\left\| \sum_{r=1}^{\mathrm{R}} t_r \mathbb{E}(Z_{r,m}B_m) - q \right\|_2 - \epsilon \right) - \nu^\top t. \tag{G.2.7}$$

We denote by $\mathcal{T}_\varepsilon$ and $\mathcal{M}_\varepsilon$ the sets of t and μ satisfying the KKT conditions, respectively.

Note that $\mathcal{M}_\varepsilon$ is a singleton. Further, we impose linear independence constraint qualification on (G.2.5) so as to ensure that $\mathcal{T}_\epsilon$ is also a singleton. Note also that (G.2.5) allows us to estimate an outer set of Γ_I and, therefore, makes our confidence intervals slightly conservative. Lastly, note that (G.2.6) is still relatively easy to calculate because it is a convex quadratic program.

We denote the unique elements of $\mathcal{T}_\varepsilon$ and $\mathcal{M}_\varepsilon$ by t_ε and μ_ε, respectively. If $\{\epsilon_{\mathrm{M}}\}_{\mathrm{M}\in\mathbb{N}}$ is a sequence of penalty terms tending to zero at a speed lower than $\sqrt{\mathrm{M}}$, then the limits of $t_{\varepsilon_{\mathrm{M}}}$

and $\mu_{\varepsilon_{\mathrm{M}}}$ are also unique. We denote such unique limits by t^* and μ^*, respectively.

Under the regularity assumptions above, it holds that

$$\sqrt{\mathrm{M}}\left(\widehat{\delta}_{\varepsilon_{\mathrm{M}}}(q;\Gamma_I)-\delta(q;\Gamma_I)\right)\xrightarrow[\mathrm{M}\to\infty]{d}\sum_{r=1}^{\mathrm{R}}t_r^*W_r+\sum_{j=1}^{\mathrm{N}+1}\mu_j^*{t^*}^\top X_j.$$

In particular, the limiting random variable is a normal random variable, whose variance can estimated from the data.

Therefore, one can derive a 95% confidence interval for any k-th component, γ_k, of $\gamma\in\Gamma_I$ by implementing the following procedure:

1. Take $q=(0,\dots,0,1,0,\dots,0)$, where 1 is in correspondence of the k-th component of q.

2. Take $\varepsilon_{\mathrm{M}}\equiv\log\mathrm{M}$. Compute $\hat{\delta}_{\varepsilon_{\mathrm{M}}}(q;\Gamma_I)$.

3. Compute the standard deviation of $\sum_{r=1}^{\mathrm{R}}t_{r,\epsilon_{\mathrm{M}}}W_r+\sum_{j=1}^{\mathrm{N}+1}\mu_{j,\epsilon_{\mathrm{M}}}t_{\epsilon_{\mathrm{M}}}^\top X_j$. Denote such standard deviation by σ_k^u.

4. Repeat Steps 1-3 with $-q$. Denote the standard deviation from Step 3 by σ_k^u.

5. Let z_α be the α-quantile of the standard normal distribution. Then,

$$[-\hat{\delta}_{\varepsilon_{\mathrm{M}}}(-q;\Gamma_I)-z_{1-\alpha/2}\sigma_k^l,\quad \hat{\delta}_{\varepsilon_{\mathrm{M}}}(q;\Gamma_I)+z_{1-\alpha/2}\sigma_k^u],$$

   is a confidence interval for γ_k with limiting coverage probability $1-\alpha$.

Note that, following Stoye (2009), we can adapt the above procedure to get uniformity with respect to the diameter of the identified set. Lastly, observe that in the previous steps we have assumed that θ_0 is known by the researcher. In practice, θ_0 is estimated and we should account for the resulting sampling uncertainty. θ_0 enters only A_m. Therefore, under the usual smoothness assumption on the behaviour of the function A_m in the neighbourhood of θ_0, we just need to modify the asymptotic distribution of W_r. Specifically, a standard Taylor expansion around θ_0 allows us to incorporate the impact of the sampling uncertainty induced by the estimation of θ_0 in the variance of W_r.

G.3 Computing the first-stage moment inequalities

We provide some directions on how to compute $\Delta\Pi_{(+ab),f}$ and $\Delta\overline{\text{FC}}_{(+ab),f}$ entering the first-stage moment inequalities in (2.16). First, we update the systematic fixed costs by adding $G_{ab,f} = 1$. Second, we update the list of products offered by firm f, by adding nonstop flights between cities a and b. Further, if a is one of firm f's hubs, then we add one-stop flights via a between b and all cities d such that $G_{da,f} = 1$. Similarly, if b is one of firm f's hubs, we add one-stop flights via b between a and all cities d such that $G_{db,f} = 1$.[4] Third, we update the matrices of product covariates by adding the demand and marginal cost shifters of the new products. Fourth, for each updated market, we randomly draw 500 vectors from a normal distribution with mean and variance equal to the empirical mean and variance of the vector of second-stage shocks that have been computed via BLP inversion. For each of these draws, we iterate on the F.O.C.s in (2.6) to find the new prices and market shares and we compute the second-stage variable profits. We have decided to use the F.O.C.s in (2.6) as a contraction mapping. While we do not formally prove that they are indeed a contraction mapping, we have found that the resulting price vector does not change when using different starting values and that the mapping converges in all the considered cases. We average across draws and get the simulated expected second-stage variable profits. Lastly, we compute $\Delta\Pi_{(+ab),f}$ and $\Delta\overline{\text{FC}}_{(+ab),f}$ as the difference between the expected second-stage profits minus the systematic fixed costs in the factual scenario and the expected second-stage variable profits minus the systematic fixed costs in the counterfactual scenario. An analogous algorithm is developed to compute $\Delta\Pi_{(-ab),f}$ and $\Delta\overline{\text{FC}}_{(-ab),f}$.

G.4 First-stage instruments

Table G.4.1 reports the instruments used to construct the first-stage moment inequalities. The first section of Table G.4.1 lists the instruments for the class of deviations "$(-ab)$". The second section of Table G.4.1 lists the instruments for the class of deviations "$(+ab)$". For example, with regards to American Airlines, we consider 3 instruments for the class of deviations "$(-ab)$":

(1) $Z_{(-ab),\text{AA}} = 1$ if cities a or b are hubs for American Airlines and market $\{a, b\}$ has a size greater than 6 millions.

(2) $Z_{(-ab),\text{AA}} = 1$ if cities a or b are hubs for American Airlines and are not historically classified as having a poor on-time performance of flights, according to the U.S. Department of Transportation.[5]

[4] We do not add a one-stop flights when the resulting itinerary is unrealistic, such as a flight from Seattle to Denver via Miami. Apart from these extreme cases, we assume that the firm will offer all possible one-stop flights.

[5] This can be found at `https://www.transtats.bts.gov/DL_SelectFields.asp?gnoyr_VQ=FGJ&QO_`

(3) $Z_{(-ab),\mathrm{AA}} = 1$ if cities a and b are not hubs for any airlines.

American Airlines will tend to always offer direct flights in the above markets, plausibly unrelated to the fixed cost shocks, due to the expected very high profitability. Still with regards to American Airlines, we consider 2 instruments for the class of deviations "$(+ab)$":

(4) $Z_{(+ab),\mathrm{AA}} = 1$ if cities a and b are hubs for the competitors and not for American Airlines.[6]

(5) $Z_{(+ab),\mathrm{AA}} = 1$ cities a or b are hubs for American Airlines and market $\{a, b\}$ has a size smaller than 3 millions.

American Airlines will tend to never offer direct flights in the above markets, plausibly unrelated to the fixed cost shocks, due to the expected very low profitability.

We construct similar instruments for the other airlines. Note that instruments (1) and (4) are based on lower and upper bounds for the market size that are homogeneous across airlines. The only exception is instrument (4) for AA, where we consider as upper bound 3 millions, while for the other carriers we take 1.5 millions. We do so in order to take account of the different observed segment choice patterns of AA.

Table G.4.1: First-stage instruments.

Markets that are served with direct flights
American: (1) hub, size $>$ 6 million; (2) hub, no historical constraints; (3) non-hub, no other firm has hub
Delta: (1) hub, size $>$ 6 million; (2) hub, no historical constraints; (3) non-hub, no other firm has hub
United: (1) hub, size $>$ 6 million; (2) hub, no historical constraints; (3) non-hub, no other firm has hub
US Airways: (1) hub, size $>$ 6 million; (2) hub, no historical constraints; (3) non-hub, no other firm has hub
Southwest: (1) hub, size $>$ 6 million; (2) hub, no historical constraints; (3) non-hub, no other firm has hub
Markets that are not served with direct flights
American: (4) hub, size $<$ 3 million; (5) non-hub, other firm has hub
Delta: (4) hub, size $<$ 1.5 million; (5) non-hub, other firm has hub
United: (4) hub, size $<$ 1.5 million; (5) non-hub, other firm has hub
US Airways: (4) hub, size $<$ 1.5 million; (5) non-hub, other firm has hub
Southwest: (4) hub, size $<$ 1.5 million; (5) non-hub, other firm has hub

`fu146_anzr=b0-gvzr`.

[6] One may wonder whether we should expect the fixed cost shocks of entering non-hub markets to be generally higher because the hub airlines may inhibit potential competitors' abilities to obtain gates, slots, and other facilities necessary for entry or expansion. We do not view this as a systematic tendency taking place at each hub airport. Further, the fact that we do not distinguish between airports in the same city lessens any concern of this type.

Appendix H

Other tables and figures

Table H.0.1: Mergers and bankruptcies.

Mergers
American Airlines + Trans World Airlines (2001)
US Airways + American West (2005)
Delta Airlines + Northwest Airlines (2008)
United Airlines + Continental Airlines (2010)
Southwest Airlines + AirTran (2010)
American Airlines + US Airways (2013)
Bankruptcies
US Airways (2002-2003)
United Airlines (2002-2006)
US Airways (2004-2005)
Northwest Airlines (2005-2007)
Delta Airlines (2005-2007)
American Airlines (2011-2013)

In the first section of Table H.0.1, we report the mergers between the major carriers after 2001. As a result of these mergers, the number of legacy carriers has dropped from 11 in 2001 to 4 nowadays. In the second section of Table H.0.1, we report the airlines which have been under Chapter 11 bankruptcy after 2001. All such bankruptcy events were resolved with a restructuring or a merger.

Figure H.0.1 represents the network of markets served by American Airlines, before the merger with US Airways. The nodes of the networks are the cities. There is a link between two nodes if American Airlines offers direct flights between those two cities. The red points represents the hubs of American Airlines (Dallas, New York City, Los Angeles, Miami, and Chicago).

Figure H.0.1: American Airlines' network.

Table H.0.2: Hubs.

AA	DL	UA	US	WN
Dallas	Atlanta	Washington DC	Charlotte	Washington DC
New York	Cincinnati	Denver	Washington DC	Denver
Los Angeles	Detroit	Houston	Philadelphia	Houston
Miami	New York	New York	Phoenix	Las Vegas
Chicago	Memphis	Los Angeles		Chicago
	Minneapolis-Saint Paul	Chicago		Phoenix
	Salt Lake City	San Francisco		

Table H.0.2 lists the hubs of the legacy carriers and the focus cities of Southwest Airlines.

Chapter 3

Barriers to adoption of real-time electricity pricing: Evidence from New Zealand

WITH CHARLES PÉBEREAU

3.1 Introduction

Economic theory predicts that introducing real-time electricity pricing (RTP) in an economy with rational and perfectly informed agents will lead the retail market to gradually unravel. The consumers with the consumption profiles least costly to serve self-select into RTP first, increasing the average cost of serving the other consumers. As a result, retailers increase their rates, and a new set of consumers finds it profitable to switch to RTP. This spiral of higher rates more switching goes on and until a significant share of consumers are on an RTP plan [1]. This scenario did not occur in New Zealand. More than seven years after the introduction of RTP, the share of residential consumers on this tariff remains below 1.25%.

In this paper, we document the introduction of RTP in New Zealand to bring new insights about barriers to adoption. Specifically, we investigate consumer behavior to identify why the New Zealand retail market did not unravel. Our findings suggest that inexperienced consumers - new and prospective adopters - overreact to contemporaneous spot prices but that, over time, consumers who have adopted RTP learn to consider long-term outcomes. In particular, we find that a crisis on the spot market led many consumers who had recently adopted RTP to switch to other tariffs. This phenomenon may have interfered with the unraveling process. The significant effect of learning is a sign of imperfect information: consumers lack information, have bounded rationality, or both. An essential remedy to this barrier is thus to provide consumers with relevant information: ex-ante, to help them in their adoption decision, and ex-post, to help them become familiar with varying spot prices and consider the long-run payoffs rather than focus on the immediate effect of spot prices.

Consumers on RTP face the time-varying marginal cost of electricity, while those on standard tariffs face a constant one.[2] Because the demand of residential consumers has historically been inelastic, the prospect of large efficiency gains stemming from increased demand response has led many economists and policy-makers to advocate for RTP.[3] In Europe, for instance, the Electricity Directive[4] aims at fostering the roll-out of smart meters and adoption of RTP for residential consumers. The literature has identified several barriers to the widespread adoption of RTP. Joskow and Wolfram (2012) review some of them and identify the fear of significant redistribution as the biggest barrier. However, the case of New Zealand shows that take-up can fail even before questions of redistribution arise. Our paper contributes to this literature by identifying imperfect information as a barrier to adoption.

[1]See for instance Borenstein (2005b).

[2]For simplicity, we refer here to flat-rates plans, which are widely popular. There exist dynamic electricity tariffs other than RTP, such as Time-of-Use or Critical Peak Pricing.

[3]See Borenstein (2005a) and Joskow and Wolfram (2012) among others.

[4]Directive (EU) 2019/944, June 2019

We study consumer behavior regarding real-time pricing in New Zealand to explain why the retail market did not unravel. In particular, we focus on the following questions: First, what affects consumer decisions to switch to real-time pricing? Second, how do consumers react to a crisis on the spot market? We answer these questions using a unique data set in which we observe all electricity retailer switches by residential consumers in New Zealand from 2014 to 2018. We further observe monthly electricity consumption, half-hourly spot prices, and detailed census data.

We find that contemporaneous electricity spot prices significantly affect consumer decisions to adopt real-time pricing. We build several price definitions that consumers may consider and show that contemporaneous spot prices explain their adoption decisions best. Furthermore, we find that, during a crisis on the spot market, most consumers choose to forego adoption rather than postponing it. While we do not observe how often consumers consider adopting RTP, this suggests that they do at most once and, thus, that they do not strategically time adoption decisions.

Following the adoption of RTP, nearly no consumers left RTP until a crisis on the electricity spot market during the winter of 2017 (hereafter referred to as the crisis). We find that during this crisis, the share of consumers switching to another tariff decreased with the time spent on RTP before the crisis. Furthermore, among consumers who discarded RTP during the crisis, the share of those switching back after the crisis increases with the time spent on RTP before the crisis. Exploiting temporal variations in the roll-out of real-time pricing across different cities, we rule out that this correlation is due to selection effects. While we document selection dynamics at adoption, this finding suggests that post-adoption learning overshadowed pre-adoption idiosyncrasies.

Finally, the demand response of consumers who stayed on RTP during the crisis decreased with their experience with the tariff. In particular, consumers with the shortest experience with RTP conserved electricity relative to consumers on other tariffs, while those with the highest experience increased their electricity consumption. This finding is at odds with selection dynamics because, intuitively, we would expect the most price-responsive consumers to adopt RTP first. This qualitative argument suggests that the correlation between experience and demand response is causal.

Together, these findings suggest that inexperienced consumers - new and prospective adopters - overreact to contemporaneous spot prices. Over time, however, consumers who have adopted RTP put less weight on immediate outcomes. This interpretation is coherent with the fact consumers on RTP have access to their savings computed relatively to their previous tariff. Thus, more experienced consumers have access to estimates of their payoffs under RTP over a longer horizon.

The perception of RTP of inexperienced consumers is paramount to the unraveling process. If prospective adopters do not switch to RTP or if new adopters quickly discard RTP, then retailers will not increase the rates on the other tariffs, which will block the unraveling process. Relatedly, if there is social learning, the first impressions of a consumer on real-time pricing can influence other consumers' decision to adopt the tariff. Indeed, consumers who have not yet adopted know consumers already having adopted are likely to benefit from RTP more than they do. Therefore, consumers who discard RTP signal to those who have yet to adopt that RTP is not for them either.

In the unraveling scenario, the first consumers who switch to RTP are those who benefit the most from it, and the benefits that the marginal consumer derives from adopting to RTP decreases in each new wave of switches. Subsequently, after each new wave, inexperienced consumers may be more easily discouraged by the advent of high spot prices. This problem may grow over time when considering that the few consumers who have adopted RTP are among the most educated but also, presumably, the most motivated to experiment with this new tariff.

Our results give rise to several policy implications. We find evidence that consumers rely on inadequate information to make decisions about RTP, either because they lack a better one or because they do not make use of it. This can interfere with the unraveling process if consumers who would benefit from RTP forego adoption or if adopters get a bad first impression and retract.

Estimating the benefits of switching is a complicated exercise and simplifying it could increase adoption. For instance, easing the access to smart meters' records of real-time electricity consumption and its usage on tariff comparison websites could help consumers identify ex-ante how valuable it is for them to switch to RTP. Furthermore, because inexperienced consumers overreact to contemporaneous spot prices upon considering adoption, it is important to explain how spot prices form and that long-run gains can compensate immediate losses. In particular, spot prices are seasonal, which means that, relative to standard tariffs, consumers adopting RTP should expect higher bills half of the year and lower bills the other half.

Policies helping consumers who have recently adopted RTP are crucial, too. We find that consumers with bad first impressions are more likely to discard RTP. Therefore, it would be useful to help them estimate long-term payoffs and compare them with ongoing losses, when they happen. In complement, one can provide some insurance to consumers who have recently adopted RTP. Alternatively, it may be necessary to avoid adoption during risky periods, such as during harsh winters or when the electric system is under stress. In particular, this calls for the strategic timing of large marketing campaigns.

Related literature. Our paper contributes to several branches of the literature. First, our paper relates to the large literature on consumer tariff choices in retail electricity markets (Crampes and Waddams, 2017; Hortaçsu, Madanizadeh, and Puller, 2017; Dressler and Weiergraber, 2019; Fowlie, Wolfram, Spurlock, Todd, Baylis, and Cappers, 2020; Ito, Ida, and Tanaka, 2021). The main focus of this literature is consumer inertia. Its goal is to identify why consumers do not switch tariff or retailer more often. To the best of our knowledge, our paper is the first to study specifically consumer decision to adopt real-time pricing and to keep or discard it. We identify overreaction to contemporaneous spot prices as an important barrier to adoption and follow-on retention of consumers on real-time pricing. This barrier is specific to real-time pricing because it is the only tariff which rates vary in real-time with spot prices.

Second, our paper relates to the literature studying demand response with dynamic electricity tariffs (Ito, 2014; Allcott, 2011; Fowlie et al., 2020; Fabra, Rapson, Reguant, and Wang, 2021; Ito et al., 2021). We cannot study real-time electricity consumption but, instead, focus on monthly electricity consumption during a crisis on the spot when spot prices were uniformly high - and therefore intertemporal substitution pointless. We show that, among consumers who remained on RTP during the crisis, demand response was a decreasing function of experience.

Third, our paper relates to the literature studying present biases and the role of first impressions on the decisions of economic agents (Busse, Pope, Pope, and Silva-Risso, 2015; Lamp, 2018; Hirshleifer, Lourie, Ruchti, and Truong, 2020). We show that contemporaneous spot prices significantly affect adoption decisions and that bad first impressions are an important driver of consumer attrition.

Finally, our paper relates to the literature studying consumer behavior and learning. One part of this literature studies consumer reactions to unexpected events such as bill shocks (Grubb and Osborne, 2015; Grubb, 2015) or unemployment (Malmendier and Shen, 2019). Another part of this literature focuses on how consumers make forecasts about future outcomes (Anderson, Kellogg, and Sallee, 2013; Malmendier and Nagel, 2016) and how they learn Miravete (2003). In the case of real-time electricity pricing we show that while inexperienced consumers overreact to contemporaneous spot prices, over time, consumers focus less on immediate outcomes.

The paper proceeds as follows. Section 3.2 describes our data and the electricity market in New Zealand. Sections 3.3 and 3.4 present our main results. Section 3.5 discusses policy implications and Section 3.6 concludes.

3.2 Data and Industry Background

3.2.1 The electricity sector in New Zealand

Historically, the generation, transport and distribution of electricity in New Zealand was publicly and centrally managed. Starting in 1987, the state-owned monopolies were split up and sold and, in 1996, a decentralized wholesale spot electricity market was established. Competition was established in generation and retailing, while transport and distribution were regulated. In July 2017 there were 40 retail brands delivered by 29 parent companies[5]. Yet, the market remains dominated by the "Big 5" (Genesis Energy, Contact Energy, Mercury NZ, Meridian Energy, Trust Power) who are vertically integrated. In 2017, they have a collective market share of around 89% in retail for residential consumers and collectively produced more than 90% of electricity in New Zealand.

While electricity is a homogeneous product, retailers can different themselves in several ways, such as bundling with other services like gas, customer service and the tariff structure. Traditionally, consumers could only have flat tariffs: two-part tariffs with known fixed and variable components. Special electricity meters allowed time-of-use tariffs with different day and night rates but flat tariffs are still dominant nowadays. New Zealand has been a world leader in the deployment of smart electricity that allow consumption to be measured in real time[6]. The roll-out of smart meters has allowed new electricity tariffs to emerge. For instance, Electric Kiwi offers two-part tariffs with one hour of free consumption per day. Other retailers such as Paua to the People and Flick Electric offer real-time pricing tariffs under which consumers pay a price indexed on the electricity spot market which clears every half-hour. Table 3.2.1 provides an example of the tariff offered by Flick Electric in 2017 in Wellington and compares it to a flat tariff offed by Contact Energy, one of the Big 5 retailers.

Table 3.2.1: Examples of electricity tariffs offered in Wellington in 2017. Note that the average spot price in the past ten years was about 7 c$/kWh and therefore, under real-time pricing, the average marginal cost of electricity is 16 c$/kWh.

Tariff Portion	Flick Electric	Contact Energy
Fix (c$/day)	177	211
Variable (c$/kWh)	9 + spot	17.6

3.2.2 Data

We make use of a unique dataset containing all occurrences of consumers switching retailers between January 2013 and June 2018. These switches are recorded at the installation control

[5]See https://www.ea.govt.nz/about-us/media-and-publications/market-commentary/events/prerequisites-for-a-competitive-retail-market/

[6]In 2017, 72% of electricity meters in New Zealand are smarty meters, the highest rate in the world.

point (ICP)-level, which is a unique identifier of an electricity meter. We observe the previous retailer, the new retailer the consumer is switching to, as well as the initiated date and the date at which the switch happened. Furthermore, we observe whether the switch was related to the household moving into the accommodation or if it occurred while he was already living there[7].

We also observe the census tract in which an ICP is located, which allows us to merge the switching data to census data from 2013. We use information on income, education, and age levels at the census-tract level (a census tract usually contains between 50 and 80 households). We have yearly and monthly electricity consumption data at the consumer-level (ie. at the ICP-level). Given that we do not observe daily variation patterns, we cannot make use of our consumption data to study how consumers on real-time pricing contracts respond to price signals.

Also, we collect publicly available wholesale price data for each network reporting region at the half-hourly level[8]. We use these half-hourly price data to compute the price faced by Flick customers. We also use the price data to compute average spot prices over different time horizons.

Finally, we have information on the history of long-term tariffs offered by each retailer. In the case of Flick, we observe the base fixed and variable fees, from which we can compute the half-hourly consumer price with the help of our spot price data.

Table 3.2.2 shows consumer characteristics of consumers in Wellington, recorded on the census tract-level. The first row shows electricity consumption in 2015, a proxy for how much the given household consumes. The next rows show the median income (in NZ$ 1,000), median age, as well as the share of households where the household head holds a higher education degree and where the household head works in a white-collar job. We provide the information for all households, those switching tariffs in the second half of 2014 and 2016, and those switching to Flick in the second half of 2014 and 2016.

3.2.3 Real-time pricing in New Zealand

We provide a brief overview of real-time pricing tariffs in New Zealand, summarized in Table 3.2.3. While several retailers offer real-time pricing, most consumers adopting this tariff con-

[7]We only observe those switches occurring due to moving where the retailer chosen by the new occupant is not the same as the retailer of the previous occupant.

[8]New Zealand is split into different network reporting regions (NRRs). Auckland, Wellington and Christchurch largely coincide with different network reporting regions, allowing us to observe the level of spot price faced by consumers on real-time pricing in those cities.

Table 3.2.2: Comparison of the median household, the median household switching retailer and the median household adopting real-time pricing - in Wellington.

	All ICPs		Switchers				RTP adopters			
			2014		2016		2014		2016	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Consumption (kWh/yr)	7.25	(3.8)	7	(3.8)	7.80	(4)	9.06	(4)	8.21	(3.8)
Income (NZ$/yr)	68.96	(29.6)	85	(31.6)	86.66	(31.7)	99.74	(29.2)	92.68	(30.7)
Age	39.55	(11.2)	36	(7.9)	36.76	(7.8)	37.15	(6.9)	36.47	(7.6)
Education (%)	18.62	(12.2)	28	(15.7)	28.98	(15.7)	36.66	(16.4)	32.15	(15.3)
Work (%)	54.32	(46.9)	50	(18.9)	48.90	(19.6)	55.32	(15.1)	52.20	(18.9)

tracted with Flick Electric[9]. Therefore, in the rest of the paper, we focus exclusively on Flick Electric for real-time pricing tariffs.

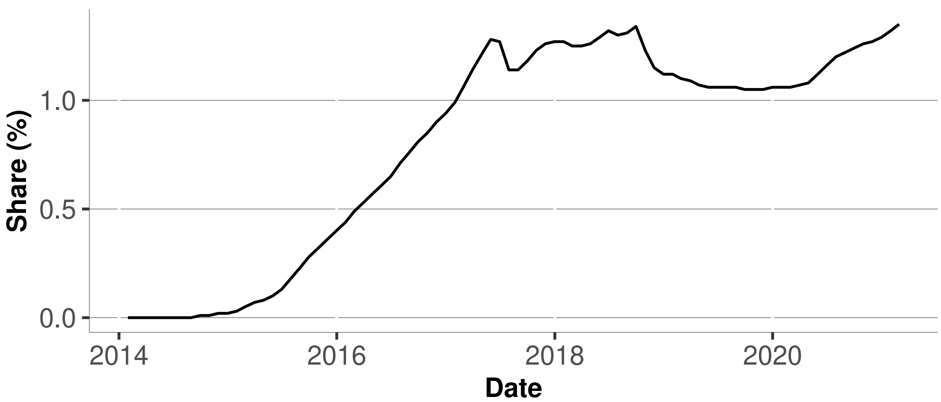


Figure 3.2.1: Evolution of the share of households on real-time electricity pricing.

Flick Electric entered the retail electricity market for the first time at the end of 2013 in Wellington and then gradually entered other cities. Figure 3.2.1 shows that its market share initially grew quickly. For instance, during the first semester of 2017, 38.1 % of households switching electricity retailers in Christchurch chose Flick Electric. At the end of May 2017, Flick Electric's market share in New Zealand was 1.28% - or 23,057 households - with large heterogeneity across cities. While in Auckland less than 1% of households chose real-time pricing, in Christchurch they were 4.46%. Flick Electric's growth stalled in June 2017 - the start of what we refer to as Winter 2017 crisis - and its market share has oscillated around 1% until (at least) the end 2020.

The Winter 2017 crisis. The Winter 2017 crisis refers to the period of high spot prices that occurred during several weeks in June, July and August 2017. This crisis was due to low

[9]In June 2017, Paua to the People had less then 1,000 customers, while Flick Electric had more than 23,000.

Table 3.2.3: Overview real-time pricing tariffs in New Zealand.

	Number of households	Market Share in May 2017	Date of Entry
New Zealand	23057	1.28	Nov 2013
Wellington	5502	3.88	Nov 2013
Auckland	5214	0.90	Jun 2014
Christchurch	6981	4.46	Sep 2015

hydro inflows coupled with high electricity demand [10]. Because in New Zealand about 60% of electricity comes from hydro generation, shortage of water meant scarce electricity generation which led to spot price increases. As illustrated on Figure 3.2.2 spot prices increased two- to three-fold compared to the previous winters. What is notable is that previous winters saw sot prices peak twice a day, once in the morning and once in the evening. To the contrary, spot prices during the winter of 2017 stay very high throughout the day and only fall during the night. However, the level of night-time spot prices in 2017 is almost as high as the peaks during previous winters. We do not have data about consumer bills. However, Flick Electric published information related to the Winter 2017 crisis on its website [11]. In particular they compare consumers savings, calculated as the difference between bills since adopting real-time pricing to what consumer bills would have been under their previous tariff. Flick reports that while consumers on real-time pricing saved, on average, about 479NZ$ in 2017 compared to their previous tariff, they made a loss of 80NZ$ alone from mid-June to mid-July during the crisis [12]. Both the price patterns as well as Flick's analysis of savings suggest that consumers faced heavy losses during the crisis and scope to contain them was limited.

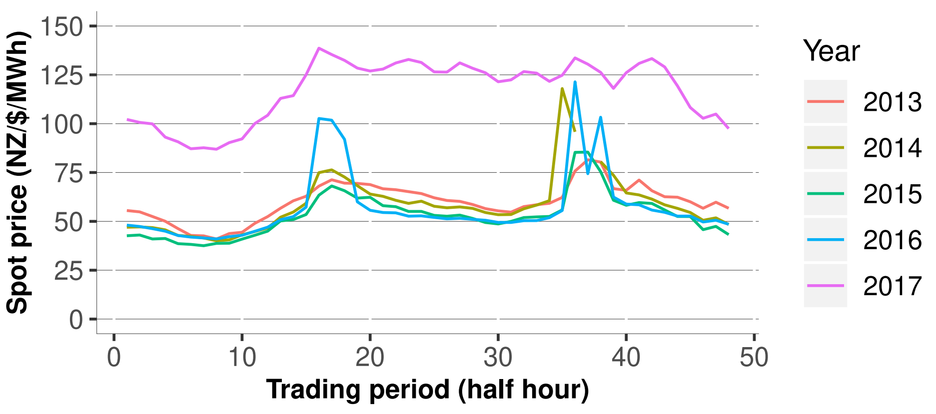


Figure 3.2.2: Average half-hourly spot price in the winters of 2015, 2016 and 2017.

[10] See Electricity Authority (2018)

[11] Available at https://news.flickelectric.co.nz/2017/08/01/the-inside-info-how-we-calculate-your-savings/

[12] For comparison, the average yearly consumer bill is around 2,200 NZ$/year

3.3 The role of spot prices in adoption decisions

In this section we study consumer adoption decisions. In particular, we investigate what prices consumers consider when deciding whether or not to switch to a real-time pricing contract [13].

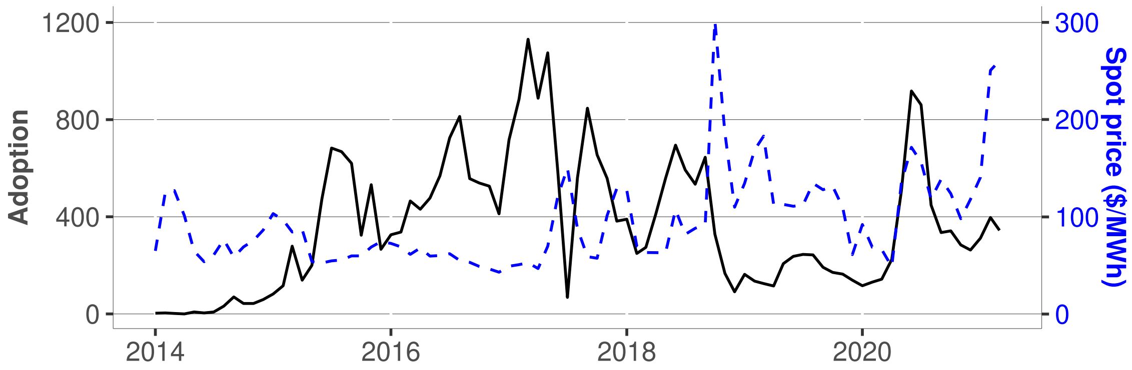


Figure 3.3.1: History of monthly RTP adoption and average spot prices.

We first present some graphical evidence of the relationship between spot prices and RTP adoption. Figure 3.3.1 plots the history of prices and share of switchers choosing RTP for Wellington, the capital of New Zealand and the city where real-time pricing was available first. We can see that drops in spot prices tend to be followed by an increase in the share of switchers choosing RTP and that increasing spot prices tend to be followed by a decrease in the share of switchers adopting RTP. For instance, in early 2015, a drop in spot prices is followed by a surge in adoption. There are two other events in 2017 and 2018, where a surge in spot prices is followed by a drop in RTP adoption rates. In each case there is a lag between the change in prices and the change in adoption, which suggests that most households do not anticipate these price surges.

At the individual level, Figure 3.3.2 plots the share of households switching retailers who choose RTP as a function of past average spot prices in the four weeks preceding the switch. The plot suggests that consumers are price-sensitive. Their probability to choose RTP upon switching drops slightly above 10% when spot prices are in the range 30-40 $/MWh to less than 5% when prices exceed 90 $/MWh.

[13] A switcher is defined as a electricity connection point ("household") that chooses to change electricity retailers.

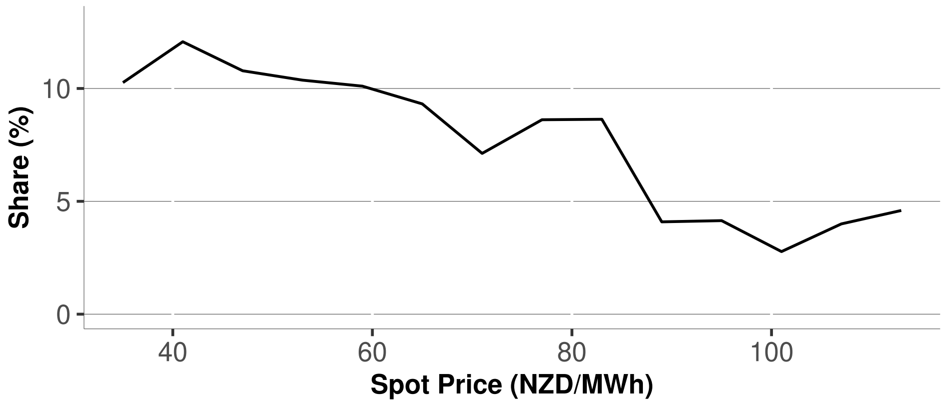


Figure 3.3.2: Share of consumers switching retailers adopting RTP as a function of average spot prices in the 4 weeks preceeding the switch - in Wellington.

3.3.1 Price definitions

In the next step, we investigate to what extend different prices are correlated with the decision to adopt real-time pricing. To do so, we build several price definitions that consumers may take into account in their decisions. We build these definitions over 1-week, 2-week, and 4-week time periods.

1. *Past Price*
   We define *Past Price* as the average spot price during the given time period before the switch. For instance, when we look at the 1-week horizon, the time period is the week before the week leading up to the switch.

2. *Recent Price*
   We define *Recent Price* to be the average spot price during the given time period immediately preceding the switch. When looking at the 1-week period, the time period is the week leading up to the switching decision.

3. *Future Price*
   We define *Future Price* to be the average price during the period following the switch. When we look at the one-week period, the time period is the week immediately following the switching decision.

4. *Last Year Price*
   We compute *Last Year Price* as the average spot price during the given time period one year before the adoption decision. *Last Year Price* is *Recent Price* from the year before.

5. *Future Predicted Price*
   We compute *Future Predicted Price* assuming that consumers predict prices using an AR(1) model. To do so, we use all spot price information available to us, resulting in AR(1) models with price data from July 2009 to the adoption decision. We assume consumers predict out weekly average prices, so we aggregate the price data to the week level.

3.3.2 Empirical strategy

Our goal is to see which price consumers consider when making the decision of whether or not to adopt RTP. In the analysis, we only make use of consumers who switch retailers. Doing so allows us to see how different prices affect the number of consumers choosing RTP, given the number of consumers having decided to switch. The analysis stays silent about effects on the overall number of switchers. Looking at this margin would require us to disentangle consumers who do not switch because they prefer staying with their current retailer and consumers who do not switch because they face early termination fees.[14] We are unable to do so in our data, so we focus on decisions to choose real-time pricing conditional on having decided to switch.

Aggregate switches

In the first part of the analysis, we regress the number of consumers adopting RTP on the first three price definitions. Formally, we estimate the following specifications:

$$Y_{mt} = \alpha_1 P_{m\tau,\text{past}} + \alpha_2 P_{m\tau,\text{recent}} + \alpha_3 P_{m\tau,\text{future}} + \varepsilon_{mt} \tag{3.3.1}$$

$$Y_{mt} = \alpha_1 P_{m\tau,\text{past}} + \alpha_2 P_{m\tau,\text{recent}} + \alpha_3 P_{m\tau,\text{future}} + \gamma_m + \lambda_t + \varepsilon_{mt}, \tag{3.3.2}$$

where Y_{mt} is the share of RTP adopters in a given week in market m, $P_{m\tau,\text{past}}$, $P_{m\tau,\text{recent}}$, and $P_{m\tau,\text{future}}$ are the different prices, with τ the time period considered (1 week, 2 weeks, or 4 weeks), γ_m are market fixed effects and λ_t are month-of-year fixed effects.

The results are in Table 3.3.1. The first three columns hold the results for specification 3.3.1 and the last three columns hold the results for specification 3.3.2. We can see that when using all three definitions, only *Past* and *Recent Price*, respectively, come up statistically significant. When defining prices over a one-week period, *Past Price* is most strongly correlated with the share of switchers choosing RTP. The coefficient on *Future Price* is significant as well in that case, but only at the 5% level. When defining prices over a two- or four-week period, it is *Past*

[14]Consumers wishing to cancel their electricity contract before the contract ends typically face an early termination fee, which can be up to around NZD200.

Table 3.3.1: Aggregate RTP adoption from 2014-06-01 to 2018-06-01

	Dependent variable:					
	Share of switchers adopting RTP (in pct)					
	1 Week	2 Weeks	4 Weeks	1 Week	2 Weeks	4 Weeks
	(1)	(2)	(3)	(4)	(5)	(6)
Past Price	−0.692***	−0.619**	−0.558**	−0.515**	−0.649**	−0.491*
	(0.192)	(0.241)	(0.230)	(0.206)	(0.264)	(0.275)
Recent Price	0.029	−0.315	−0.542***	−0.036	−0.524**	−1.150***
	(0.243)	(0.254)	(0.196)	(0.178)	(0.213)	(0.343)
Future Price	−0.262	−0.099	−0.152	−0.335**	−0.179	0.246
	(0.209)	(0.206)	(0.218)	(0.164)	(0.188)	(0.300)
NRR FE?	No	No	No	Yes	Yes	Yes
Year- Month FE?	No	No	No	Yes	Yes	Yes
Observations	617	617	617	617	617	617
R^2	0.068	0.076	0.087	0.617	0.621	0.623
Adjusted R^2	0.063	0.072	0.082	0.581	0.585	0.588
Residual Std. Error	8.548	8.509	8.462	5.715	5.686	5.671
F Statistic	14.877***	16.910***	19.380***	17.139***	17.416***	17.572***

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

Price (two weeks) and *Recent Price* (two and four weeks) that come up statistically significant. To give an idea of the magnitudes, the coefficient on *Recent Price* in column 6 suggest that when spot prices increase by 1 $ct\$/kWh$, the RTP adoption rate drops by -1 percentage points. In our sample, the average spot price is $7.13ct\$/kWh$ and the average weekly RTP adoption rate is 8.7%, meaning the result is economically significant.

Individual switches

We next move to the household level and analyze the link between individual decisions of adopting RTP and our price definitions. We employ a logit model, where we regress the individual decision to RTP pricing on the different price definitions, controls, and fixed effects. We run two specifications: one with only one price definition at a time and one with all three together. Using only one price definitions at a time allows us to circumvent potential issues related to the high correlation between the different price definitions. Also, using one price at a time allows us to formally test which model explains the data best. Formally, our specifications can be written as

$$Y_{imt} = \alpha_d P_{m\tau,d} + X_{imt}\beta + \gamma_m + \lambda_t + \varepsilon_{imt} \tag{3.3.3}$$

$$Y_{imt} = \alpha_1 P_{m\tau,\text{past}} + \alpha_2 P_{m\tau,\text{recent}} + \alpha_3 P_{m\tau,\text{future}} + X_{imt}\beta + \gamma_m + \lambda_t + \varepsilon_{imt}, \tag{3.3.4}$$

where Y_{imt} is now an indicator if consumer i in market m at date t decides to adopt RTP (conditional on switching), $P_{m\tau,d}$ is the given price considered with $d \in \{\text{past, recent, future}\}$, X_{imt} holds control variables and ε_{imt} is assumed to follow a logistic distribution. In X_{imt}, we control for total household consumption as well as consumption differences between winter and summer, the origin retailer, the Winter 2017 crisis, and census-level median household income, age, and work- and education status. Implicitly, we assume that a consumer who has made the decision to switch retailers adopts RTP if and only if her utility from adopting RTP is larger than from not doing so (and choosing a conventional tariff). Consumer utility is a latent variable. We only observe the switching decision that is an indicator of whether the consumer's utility from adopting RTP is larger than her utility of choosing a conventional tariff.

We focus on the results based on the two-week window here. The results are in Table H.3.5. The results using the one- and four-week window are in Tables H.1.1 and H.1.2 in Appendix H.1. The first three columns in each table hold the results for specification (3.3.3) and column 5 holds the results for specification (3.3.4). We first focus on the last column in the three tables. We see that both *Past Price* and *Recent Price* are both economically and statistically significant. *Future Price*, however, is not statistically, nor economically, significant. These results lead us to conclude that it is predominantly *Past Price* and *Recent Price* that play a role in consumer adoption of RTP, with this conclusion being robust across different time periods. To get an idea of the magnitudes, the coefficient on *Recent Price* in column 5 of Table H.3.5 means that at average value of the covariates, an increase in the spot price by one standard deviation decreases the probability of adopting real-time pricing by around 0.58 percentage points. Note that at the average value of covariates, the probability of adoption is around 8.09%.

We now turn to the results in columns 1-3. At first sight, the conclusions we can draw from these regressions seem less clear, as all prices definitions are statistically significant. However, *Future Price* is significant only at the 10%-level. Also, these regressions allow us to formally test which price definition best fits the data, using the Vuong (1989) test. The test allows us to compare models pair-wise. The results are in Table 3.3.3.

In the first row, we compare the specifications using *Past Price* and *Recent Price*, respectively. We can see that except for the four week horizon where *Recent Price* is preferred, neither definition clearly dominates. In second and third rows, we compare *Past Price* and *Recent Price* with *Future Price*. We can see that in these cases, the test never prefers *Future Price*. We view these results as a confirmation of the findings using the specifications in (3.3.3) and (3.3.4).

As a further robustness check, we run similar regressions with the last two of our price specifications. The results are in Table H.1.3 in Appendix H.1. These results confirm our findings above. Contemporaneous spot prices significantly affect consumer decision to adopt real-time electricity pricing.

Table 3.3.2: RTP-adoption, two-week period, data from 2014-06-01 to 2018-06-01

	Dependent variable:			
	Adopt RTP			
	(1)	(2)	(3)	(4)
Past Price	-0.084^{***}			-0.068^{***}
	(0.011)			(0.012)
Recent Price		-0.089^{***}		-0.075^{***}
		(0.011)		(0.012)
Future Price			-0.020^{*}	0.009
			(0.011)	(0.012)
Electricity consumption	0.037^{***}	0.037^{***}	0.037^{***}	0.037^{***}
	(0.003)	(0.003)	(0.003)	(0.003)
Seasonal consumption difference	-0.005	-0.004	-0.005	-0.004
	(0.033)	(0.033)	(0.033)	(0.033)
Income	1.593^{***}	1.606^{***}	1.623^{***}	1.586^{***}
	(0.463)	(0.463)	(0.463)	(0.463)
Age	-0.009^{***}	-0.009^{***}	-0.009^{***}	-0.009^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
Work	0.550^{***}	0.552^{***}	0.550^{***}	0.552^{***}
	(0.113)	(0.114)	(0.113)	(0.114)
Education	1.258^{***}	1.257^{***}	1.256^{***}	1.258^{***}
	(0.122)	(0.122)	(0.122)	(0.122)
NRR FE?	Yes	Yes	Yes	Yes
Year-Month FE?	Yes	Yes	Yes	Yes
Retailer Origin FE?	Yes	Yes	Yes	Yes
Observations	143,435	143,435	143,435	143,435
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 3.3.3: Outcomes of Vuong tests on different specifications

	1 week		2 weeks		4 weeks	
	Value	Preference	Value	Preference	Value	Preference
Past v. Recent	1.974	Past	-0.371	Neither	-3.395	Recent
Past v. Future	3.253	Past	3.517	Past	2.364	Past
Recent v. Future	2.170	Recent	3.420	Recent	4.596	Recent

Note:
Tests conducted at 5%-level.

3.3.3 Discussion

Do households forego or postpone adoption?

The results from the previous section suggest that households are price sensitive and, upon considering adopting RTP, react to recent and past spot prices. A natural follow-up question, then, is whether when spot prices are high, households forgo adopting RTP altogether or postpone adoption. To answer this question, we focus on the Winter 2017 crisis when spot prices surged and remained high for several weeks during June, July, and August. The crisis was due to a combination of low hydro levels and large electricity demand for heating in winter. Because spot prices more than doubled, consumers who were willing to adopt RTP and were able to postpone adoption had an interest in doing so. Looking at Figure 3.3.1, we clearly see that during the crisis, households stopped adopting RTP. However, when prices start declining in August, there is a surge of RTP adoption. This surge may be due to households who waited strategically or simply due to prices falling sharply. In order to get an idea whether and to what extend consumers strategically waited, we use our results from aggregate switching to predict the number of consumers who would have chosen RTP had prices between June 2017 up to the end of September 2017 been at the average of 2014 to 2016. We make use of the 2-week window (corresponding to column 5 in Table 3.3.1). The results are in 3.3.3. The observed share of switchers choosing RTP are in purple, whereas the counterfactual share is in red. The dotted red lines denote the 95% confidence interval.[15]

[15] We obtain the confidence interval through a bootstrapping procedure. We draw coefficients from a normal distribution with mean equal to the vector of estimates and the variance-covariance matrix equal to the estimated variance-covariance matrix coming from the regression model. We then take 100 draws, predict switching shares, and take the 2.5th and 97.5th percentiles, respectively.

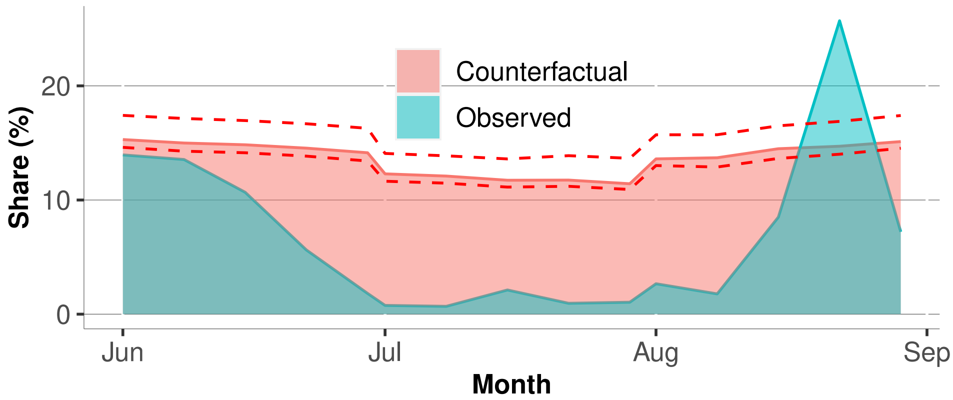


Figure 3.3.3: Share of consumers switching electricity retailer choosing RTP for the first time and electricity spot prices - biweekly in Wellington in 2017

We see that, had prices been at the average of the previous years both during as well as just after the crisis, a much larger share of consumers would have adopted RTP. At the mid-point of our prediction, 719 more consumers would have adopted (compared to 883 who did adopt over this time period). These results suggest that the overwhelming majority of consumers chose to forego adoption altogether. The fact that consumers seem to not time their adoption strategically provides further evidence that the switching decision is indeed strongly influenced by contemporaneous spot prices and not due to forward-looking behavior and strategic waiting.

Who adopts real-time pricing

The results of this section also shed light on who adopts real-time pricing. In Table H.3.5, we see that the socio-economic variables that we use as controls are also strong predictors of the decision to adopt RTP. These results confirm what we see in Table 3.2.2: On average, the consumer adopting RTP consumes more electricity, has a higher income, more education, and more likely to work in a white-collar job. These consumers are early adopters, who tend to differ from the general population quite substantially. The fact that these consumers are different from the general population has important implications for our results: It is likely that these early adopters are more interested in RTP and more sophisticated compared to the average consumer. Nevertheless, even these sophisticated early adopters draw on simple heuristics in the form of contemporaneous spot prices to make their adoption decisions. Hence it is likely that later adopters would do so as well, perhaps even more strongly so. Such a behavior violates a key condition for unraveling to occur: Consumers need to be able to self-identify as structural winners, which requires a computation of long-run savings under RTP. The fact that contemporaneous spot prices play such a large role in the adoption decision suggests consumers are not able or willing to do so.

3.4 The role of experience in consumer attrition

In this section, we study the behavior of consumers on real-time pricing during the Winter 2017 crisis. Spot prices increased significantly and remained high for several weeks which directly affected them. Because consumers could not hedge their risks on financial markets, they could only protect themselves by adjusting their electricity consumption or by switching to another tariff. We first investigate the consumer decision to discard RTP during the crisis, before moving to consumption decisions.

3.4.1 Discarding real-time pricing

In this section we focus on how and to what extent consumers reacted on the extensive margin by leaving real-time pricing during the crisis. We first present some graphical evidence. Figure 3.4.1 plots the share of consumers on real-time pricing at the start of the crisis who decide to switch to another tariff during the crisis as a function of the number of days they have spent on RTP prior to the crisis. [16] The curve is decreasing which means that consumers who have spent the most time on RTP are less likely to switch. For instance, nearly 25% of consumers who have spent less than 50 days on RTP prior to the crisis switch while only 10% of those who had spent more than 600 days on RTP do.

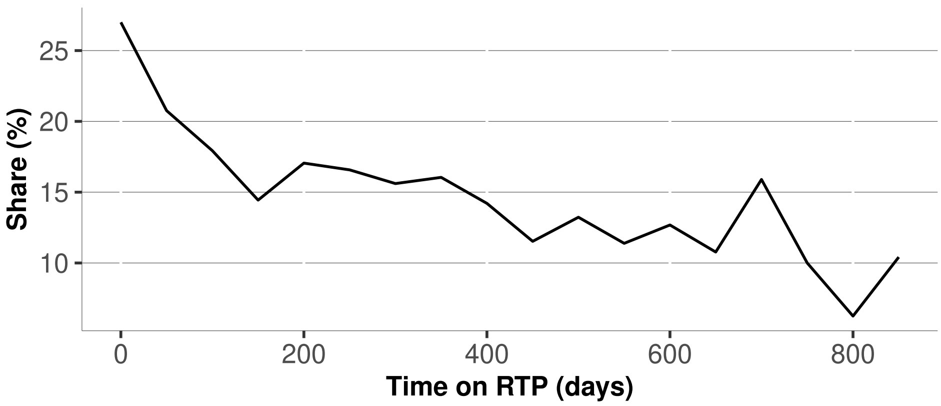


Figure 3.4.1: Share of household on RTP switching to another tariff during Winter 2017 as a function of the time they have spent on RTP.

We assume that most consumers on real-time pricing were aware of the crisis and thus had to make a conscious decision to stay with this tariff or to switch to another one. This assumption seems reasonable because information about the crisis was widely available. The crisis occurred in winter when electricity consumption usually high and consumers on real-time pricing were billed weekly. Furthermore, they received information about the crisis by their retailer and the

[16] That is, we measure the number of days between the date of adoption of real-time pricing and the beginning of the Winter 2017 crisis that we set on June 1st, 2017

event was also covered in the media. As a result, there are two reasons why consumer decisions during the crisis would be correlated with how long they have spent with real-time electricity tariffs. First, there may have been selection bias; different types of consumers adopting RTP over time. Second, after adopting real-time pricing, consumers may have changed with experience.

The economic and marketing literature on the dynamics of technology adoption documents that different categories of consumers adopt an innovation at different times. Typically, initial adopters are more motivated to try the innovation than because they value it more. Electricity is a homogeneous good, therefore the value of real-time pricing must come from the savings that certain consumers can expect compared with other tariffs [17] or from the nature of the tariff itself - such as the curiosity or the interest of facing varying and uncertain prices.

Table 3.4.1: Comparison of early and late adopters in Wellington and Christchurch

	Wellington				Christchurch			
	Early		Late		Early		Late	
	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
Consumption (kWh/yr)	9.11	(3.8)	8.35	(3.9)	10.31	(4.7)	10.00	(4.5)
Income (NZ$/yr)	98.74	(28.7)	93.71	(29.7)	75.00	(25.7)	70.84	(23.9)
Age	37.20	(7.1)	36.34	(7.3)	38.15	(8.1)	38.03	(7.8)
Education (%)	36.20	(16.7)	33.68	(15.3)	19.86	(9.7)	18.23	(9.2)
Work (%)	55.27	(15)	52.94	(17.1)	40.09	(17.5)	38.34	(19.1)

Either way, because the tariff is risky and complicated and because electricity - unlike, say, cars - is not a good for which many consumers show interest, we can expect the first adopters to be rather different than the consumers adopting later. In the last two columns of Table 3.4.1 we compare some observable characteristics of consumers adopting real-time pricing in the first semester of 2014 and 2016 in Wellington. We see that the average consumer adopting RTP in 2014 has a higher socio-economic status (higher income and work category and more educated) and higher electricity consumption than the one adopting RTP two years later. In the literature it is often argued that socio-economic status correlates with unobserved consumer traits such as myopia or risk aversion. Rich and educated consumers are thought to be more sophisticated, less risk averse, and can afford experimenting a risky contract. If this is the case, then selection bias might explain the correlation between consumer decisions and time spent on RTP.

[17] The literature on real-time electricity pricing typically distinguishes structural winners - who naturally prefer to consume electricity off-peak when spot prices are low - from structural losers. Furthermore, consumers have heterogeneous price elasticities and some can more easily adapt to varying spot prices than others See Borenstein (2005b).

Another possible explanation for the fact that consumers who spent less time on real-time pricing were more likely to leave during the crisis can be that the experience with the tariff plays a role, rather than any selection. While Table 3.4.1 presents some evidence of selection effects at the time of adoption, it must not necessarily be the case that this selection effect plays a role during the crisis.

In order to investigate the role of selection and experience in the decision to leave real-time pricing during the Winter 2017 crisis, we first regress the decision to discard real-time pricing during the crisis on observable consumer characteristics, controlling or not for their experience with the tariff. Formally, the specifications write

$$\text{Discard RTP}_{im} = X'_{im}\beta + \varepsilon_i \tag{3.4.1}$$

$$\text{Discard RTP}_{im} = \alpha \text{Time on RTP}_i + X'_{im}\beta + \varepsilon_i, \tag{3.4.2}$$

where Discard $\text{RTP}_{im} \in \{0, 1\}$ is an indicator equal to 1 if and only if consumer i in market m decides to discard RTP during the Winter 2017 crisis, Time on RTP_i is the number of months that consumer i spent on RTP prior to June 1st, 2017 - X_{im} contains control variables and ε_{im} is assumed to follow a logistic distribution.

Table 3.4.2: Discarding RTP during Winter 2017 crisis: selection vs experience

	Dependent variable:					
	Discard RTP					
	(1)	(2)	(3)	(4)	(5)	(6)
Time on RTP (months)	−0.03***		−0.03***		−0.04***	
	(0.004)		(0.005)		(0.005)	
Income (k$/yr)	0.0005	0.0004	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	−0.01***	−0.01***	−0.01**	−0.01**	−0.01**	−0.01**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Work (%)	−0.01	−0.01*	−0.01	−0.01*	−0.005	−0.01
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Education (%)	−0.003	−0.003	−0.003	−0.004	−0.003	−0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Location FE?	Yes	Yes	Yes	Yes	No	No
Previous retailer FE?	No	No	Yes	Yes	No	No
Location-on-Previous retailer FE?	No	No	No	No	Yes	Yes
Observations	8,617	8,617	8,617	8,617	8,617	8,617
Log Likelihood	−4,085.61	−4,116.60	−4,039.31	−4,068.43	−4,022.86	−4,052.50
Akaike Inf. Crit.	8,191.21	8,251.20	8,132.62	8,188.86	8,143.72	8,201.00

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

In X_{im}, we control for household yearly electricity consumption as well as consumption differences between winter and summer, the previous retailer, and census-level median household income, age, and work- and education indexes The results are in Table H.3.1. We find that the coefficients for socio-economic factors are all economically insignificant. On the other hand, the coefficient for time spent on the tariff is both statistically and economically significant. At average value of the covariates, spending 4 more months on real-time pricing decreases the probability to discard real-time pricing by 2.06 percentage points (the average probability to opt-out of 19.03%). Furthermore, controlling or not for time spend on RTP does not affect the other coefficients. These observations suggest that experience does indeed affect consumers' perception of the tariff. These observations suggest that selection dynamics do not explain consumers' decisions to discard real-time pricing during the crisis and therefore that the variable 'Time on RTP' captures the effects of experience.

For robustness, we run the same regressions and also control for a "first experience" effect for consumers who joined last with dummy, called 'Joined Last', equal to one if the consumer adopted RTP with the last cohort. The goal is to ensure that the measured effect of experience is not solely driven by the last adopters, but that experience affects all adopters. In Table H.3.2 in Appendix H.1, we test five specifications when the last cohort consists of all consumers joining during the $m \in \{1, 2, 3, 6, 9\}$ months preceding the Winter 2017 crisis. The effect is positive, statistically and economically significant, and strongest for the cohort that joined two months before the crisis started. Furthermore, controlling for this bad first impressions doesn't affect the coefficient for experience. This suggests that both first impressions and experience matters.

3.4.2 Experience and the lagged roll-out of real-time pricing

In order to get a more causal interpretation of the effect of experience, we exploit time lags in the roll-out of real-time pricing tariff across cities. The tariff was available in Christchurch only around two years after it was first introduced in Wellington. Thus, if selection bias were an important driver for switching decisions during the crisis then we would expect that two households adopting real-time pricing at the same time in Wellington and Christchurch make two different switching decisions. They have the same experience with the tariff but are two very different type of consumers- the consumer in Christchurch is an early adopter whereas the consumer in Wellington is a late mover closer to the average consumer. We rely on two key assumptions here. First, we assume that over two years, the type of consumer adopting real-time pricing changes significantly, such that the early adopter in Wellington is significantly different from the consumer adopting two years later. Table 3.4.1 suggests this is indeed the case. Second, we assume that the initial adopters in Christchurch are significantly different from the consumers in Wellington adopting real-time pricing at the same time after controlling for observable consumer characteristics. Figure 3.4.2 below shows that when RTP became available in Christchurch, 39% of all consumers in Wellington who adopted RTP before the crisis had already already adopted it.

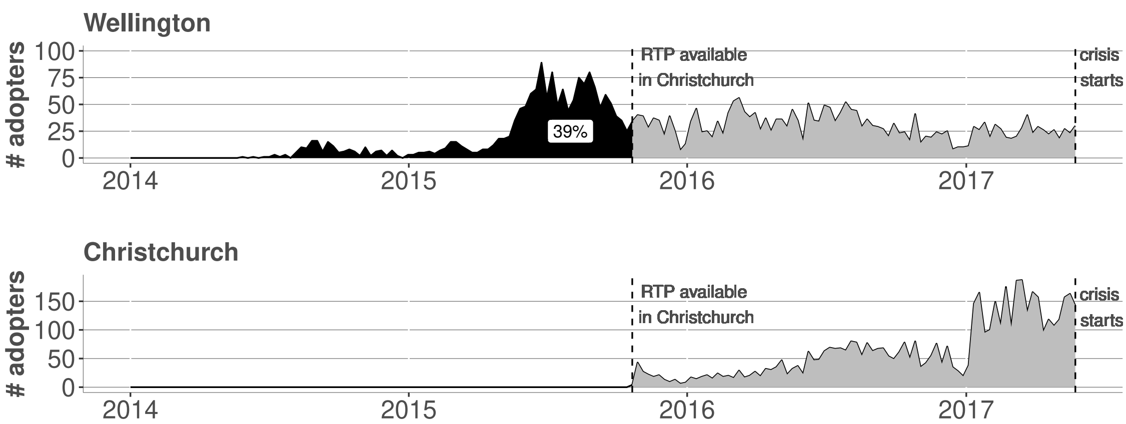


Figure 3.4.2: Number of consumers adopting real-time electricity pricing every week in Wellington (top) and Christchurch (bottom) between November 2013 and June 2017.

To see whether selection effects are present or whether it is experience that matters, we regress consumer decisions to opt-out from RTP during the crisis (Discard RTP$_{im} \in \{0,1\}$) on experience (Time on RTP$_i$), a location dummy for Christchurch, an interaction between the location dummy and control variables:

$$\text{Discard RTP}_{im} = \alpha_{exp}\text{Time on RTP}_i + \alpha_{loc}\text{Christchurch} + \gamma\text{Time on RTP}_i \times \text{Christchurch} \tag{3.4.3}$$

$$+ X'_{im}\beta + \varepsilon_i, \tag{3.4.4}$$

where ε_i is a logistic error term. Our sample is the set of households who adopted real-time pricing in Wellington and Christchurch only after the tariff was available in Christchurch, in September 2015. Therefore, the Wellington sample is truncated - we have removed the initial adopters - while the Christchurch sample is not. If selection bias were an important driver, the interaction variable γ between the location dummy and experience would be statistically significant. However, Table H.3.3 shows that in all specifications, the interaction variable is statistically - and economically - insignificant. The insignificance of γ suggests that experience, rather than selection, explains the correlation between time spent on the tariff and decision to stay or opt-out during winter 2017 crisis. Any selection effects explaining differences in the types of consumers adopting at different moments do not play an effect in these consumers' decision to leave during the Winter 2017 crisis. For robustness, we repeat the same exercise between Auckland and Christchurch. The results are in Table H.1.5 in Appendix H.1. We again find the interaction term to be insignificant as well. We conclude that consumers' perception of real-time electricity tariff is affected by their experience with the tariff. As a result, the probability that a consumer on real-time pricing switches to another tariff during the Winter 2017 crisis decreases with her experience.

Table 3.4.3: Probability of discarding RTP: Wellington vs Christchurch, only observations after RTP becomes available in Christchurch

	Dependent variable:		
	Discard RTP		
	(1)	(2)	(3)
Time on RTP (month)	-0.04^{***}	-0.04^{***}	-0.04^{***}
	(0.01)	(0.01)	(0.01)
Christchurch	0.18	0.15	0.30
	(0.14)	(0.14)	(0.34)
Time on RTP x Christchurch	−0.004	0.003	0.01
	(0.01)	(0.02)	(0.02)
Month FE?	No	Yes	No
Month-on-NRR FE?	No	No	Yes
Observations	6,142	6,142	6,142
Log Likelihood	−3,109.52	−3,098.54	−3,094.68
Akaike Inf. Crit.	6,239.04	6,239.09	6,253.36
Note:	*p<0.1; **p<0.05; ***p<0.01		

Switching back to real-time pricing. Furthermore, we document that 26.55% of households who discarded RTP during the Winter 2017 crisis switched back to RTP after prices have fallen. We show here that the probability that a consumer returns to real-time pricing after the crisis increases with their experience with the tariff. Given that only consumers with a good perception of the tariff would come back, this observation reinforces Result 2: experience affects perception. On Figure 3.4.3 we plot the share of such consumers as a function of their experience with the tariff, in Wellington. Among the consumers who discarded real-time pricing during the Winter 2017 crisis, those with less than 100 days of experience are about 15% to switch back to RTP while more than 30% of those with more than 500 days of experience do. We confirm this graphical evidence by regressing (logit) consumers decision de return to real-time pricing on their experience and control variables, see Table H.1.6 in Appendix H.1. We find that experience significantly affects the decision to return to real-time pricing, both significantly and economically. At the average of the covariates, increasing experience by 4 months increase the probability to return to RTP after the crisis by 4.8 percentage points. This result suggests that consumers who have little experience and leave may be scared by their experience during the crisis and consequently don't return. More experienced consumers leaving, on the other hand, may merely "sit out" the crisis on a contract giving them price certainty. They essentially game the system. However, the result we find, why statistically and economically significant, is rather small. Also, the overall incidence of consumers leaving and coming back is low- only around 5.52% of consumers did so during the crisis.

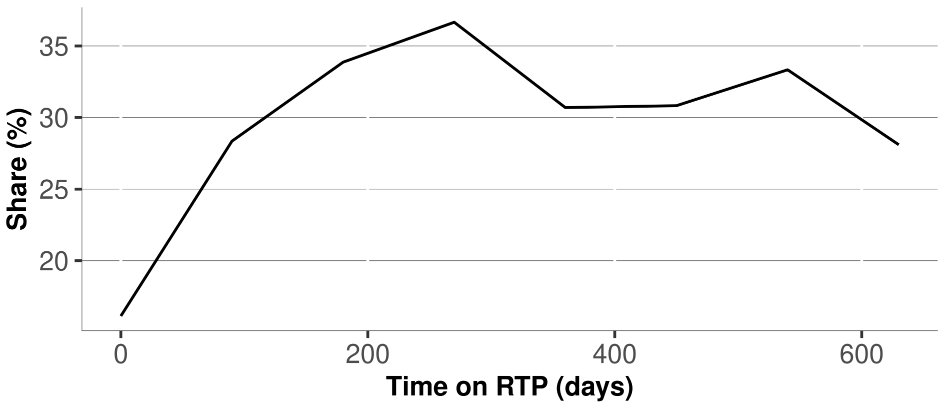


Figure 3.4.3: Share of households switching back to RTP after opting-out during the Winter 2017 crisis within three months after the end of the crisis as a function of the number of days spent on RTP prior to the crisis.

3.4.3 Adjusting electricity consumption

In this subsection we focus on consumers who remained on real-time pricing during the Winter 2017 crisis and study whether they changed their electricity consumption. Consumers can shift consumption from periods with high spot prices to periods with lower spot prices or they can reduce consumption altogether. Shifting consumption was hard to do during the crisis, as suggested by Figure 3.2.2. The figure compares the average daily pattern of spot prices during the winters of 2013 to 2017. During the winter 2017 crisis spot prices were 117 NZ$/MWh on average, or about 2.2 times larger than in winter 2015. Furthermore, spot prices vary greatly between peak and off-peak hours in winter 2016 but are rather uniform during the winter 2017 crisis. Overall, during the Winter 2017 crisis spot prices were uniformly high and therefore there is little scope and incentive for consumers to adjust their consumption throughout the day. As a result, the main strategy for consumers to lower their bills is to reduce their consumption. Unfortunately, we do not have consumption data at the trading-period (i.e. half-hourly) level to formally test this hypothesis. However, we do dispose of monthly consumption data, allowing us to test the prediction that consumers would have mainly reacted by reducing consumption overall.

Reducing consumption. In a first step, we investigate whether consumers on real-time pricing reacted fundamentally different than consumers on fixed contracts. In order to do so, we rely on a difference in differences approach: We use consumer consumption in 2014 (or 2015) and 2017. We regress individual consumption on a dummy equal to one if consumption occurs during the crisis, a dummy equal one if the consumer is on real-time pricing, an interaction between those dummies, and controls. Formally, the regression writes

$$Y_{it} = \alpha \text{Crisis}_t + \gamma \text{RTP}_i + \delta \text{Crisis}_t \times \text{RTP}_i + X'_{m_i}\beta + \varepsilon_i \tag{3.4.5}$$

Table H.1.8 in Appendix H.1 shows the results for regressions with and without controls, as well as specifications where we use winter 2015 consumption instead of winter 2014 and the mean of those two years. We see that the interaction term is negative and statistically significant when comparing consumption in the winter of 2015 to the winter of 2017 as well as in the case where we compare the mean across 2014 and 2015 with the winter of 2017. The value of the coefficient on the interaction term in column 6 implies a reduction of consumption equal to around 1.65% of the average consumption of consumers on RTP in 2017. This effect is economically significant, albeit not very large. The rather small magnitude of this effect makes sense, given that large reductions in electricity consumption across a period of three months is hard to achieve, especially when a substantial portion of electricity is consumed for heating.

The role of experience. In a next step, we investigate the relationship of experience with real-time pricing and reactions along the intensive margin. Figure 3.4.4 plots the change in consumers monthly consumption between the winters of 2014 and 2017. The black plain line corresponds to the consumption change of consumers on real-time pricing as a function of their experience with the tariff and the dashed red line shows the average change of consumers not on RTP. These consumers serve as a benchmark because they do not have spot contracts and therefore their electricity consumption is not directly affected by the crisis. We compare differences between the winter of 2014 and the winter of 2017 for two reasons: The first is that consumption patterns are seasonal, mainly because of electric heating, so two winter periods are comparable. The second is that in the winter of 2014, no consumer was on real-time pricing, so the comparison is not polluted by structural changes due to switching from a fixed to a real-time pricing tariff in between. While there are large fluctuations across experience levels, we can see a drop in consumption especially for consumers who spent less than 200 days on RTP prior to the beginning of the crisis. This drop suggests that consumers with more experience may be less responsiveness on the intensive margin than less experienced ones. This result also seems to contradict selection effects under which savvy, price-responsive consumers adopt first, as we would expect those consumers to adjust consumption more.

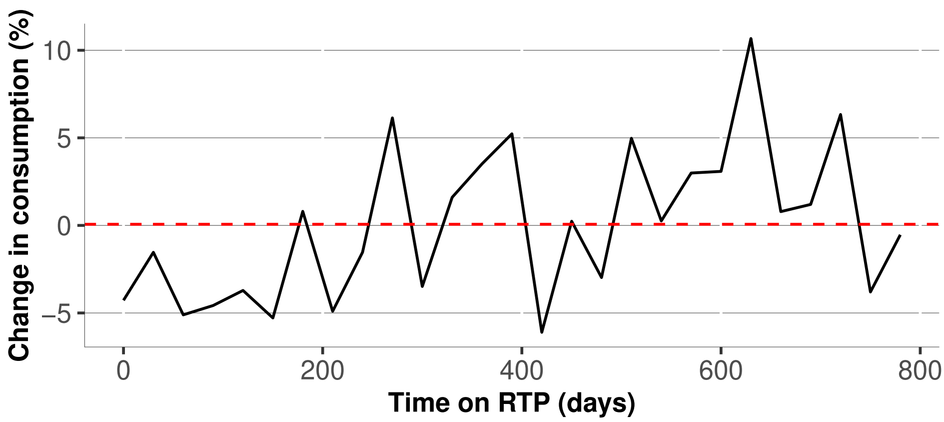


Figure 3.4.4: Change in average winter electricity consumption between 2014 and 2017 as a function of consumer experience on RTP. The red dashed line is the average change for consumers who are not on RTP.

We now turn to a more thorough investigation of the link between experience with real-tie pricing and consumption reduction during the crisis. We focus on consumers on real-time pricing at the start of the Winter 2017 crisis and who remain with this tariff during the tariff. We regress individual monthly consumption in winter 2017 on past monthly winter consumption, experience with the tariff, and controls. The regression writes

$$\text{Cons Winter } 2017_i = \alpha \text{ Time on RTP}_i + \gamma \text{ Cons Winter } 2014_i + X'_{m_i}\beta + \varepsilon_i \tag{3.4.6}$$

We run six different specifications: linear-on-linear or log-on-log, and using consumption in winter 2014, 2015 or both. The results are in Table H.3.4 in Appendix H.1. In each case the coefficient α for experience is positive and statistically significant. Economically, one additional year of experience leads to an average increase of monthly consumption of 0.7% compared with the average consumption during the winter 2017 crisis. This result suggests that consumers who have less experience with real-time pricing react more strongly along the intensive margin than more experienced consumers. There exist several potential explanations for this result. One explanation can be that consumers that have been on real-time pricing longer react less to high spot prices because they have already saved a lot by having been on real-time pricing for a while. This is especially true for consumers who adopted RTP from a Big-5 retailer, as those tend to have the highest prices. In order to see whether such an effect exists, we re-use the specifications from (3.4.6) and complement it by a dummy indicating a consumer was previously with a Big-5 retailer. Formally, we have

$$\text{Cons Winter } 2017_i = \alpha \text{ Time on RTP}_i + \delta \text{ Big-5}_i + \gamma \text{ Cons Winter } 2014_i + X'_{m_i}\beta + \varepsilon_i \tag{3.4.7}$$

The results are in Table H.1.10 in Appendix H.1. We do not find evidence that retailers from Big-5 retailers reduced consumption by less.

Familiarity bias when leaving real-time pricing

In order to decide whether to remain on real-time pricing or to switch to another tariff, rational consumers would compare their expected utility with each tariff and take into account switching costs (psychological, contractual, opportunity cost) and uncertainty regarding future spot prices and changes of tariffs. In practice, however, this is a complicated decision-making process and there is a large and growing literature showing that economic agents often use simple heuristics to make decisions. In our data, we indeed find evidence that consumers' switching decisions during the Winter 2017 crisis were not only affected by their experience with real-time pricing, but also by their previous experiences.

Table 3.4.4: Test

		Retailer Origin		
	Average	Big 5	Non-Traditional	Others
Origin of RTP-adopters (%)	NA	73.6	17.5	8.8
Switch during Winter 2017 crisis (%)	19.1	17.4	23.9	23.2
Switch to Big 5 (%)	35.4	41.7	21.8	23.8
Switch to Non-Traditional (%)	52.0	47.8	66.2	48.7
Switch to Others (%)	12.7	10.5	11.9	27.5

We split retailers into three categories and we study whether consumers switching decisions during the Winter 2017 crisis correlate with which category of retailers they were contracting with prior to adopting real-time pricing. The first category, 'Big 5', is composed of the five largest retailers. Retailers of the second category, 'Non-Traditional', offer non-traditional tariffs, such as one hour free per day, prepaid electricity, etc. Finally, we label as 'Others' the rest of the retailers - composed of both small players and subsidiaries of the Big 5. Table H.1.11 in Appendix H.1 summarizes the results. First, consumers who contracted with a Big 5 prior to adopting real-time pricing are less likely than the others to switch to another tariff during the Winter 2017 crisis. While a large majority (73.6%) of RTP-adopters were previously with a Big 5 retailer, only 17.42% of them switch to another tariff during the crisis - or about 6 percentage points less than the other consumers. Second, consumers opting-out from real-time pricing during the Winter 2017 crisis are disproportionately more likely (by about 20 percentage points) to switch to a retailer from their category of origin than other consumers. For instance, 41.7% of consumers who were with a Big 5 retailer prior to adopting real-time pricing return to a Big 5 retailer during the Winter 2017 crisis while only 21.8% and 23.8% of the other consumers do. For robustness we draw the same table but focus instead on consumers leaving real-time pricing and not returning after the crisis; that is, we do as if consumers who leave and then return to RTP had never left - see Table H.1.11 in Appendix H.1. We find the same results which confirms that, upon switching from real-time pricing, the choice of retailer is unlikely to be motivated by the perspective to come back.

We argue that these results are explained by the fact that consumers put a large weight on their personal experience when they make decisions, a phenomenon we label as familiarity bias. First, if consumers rely primarily on their experience then they are only aware of few alternatives and therefore, upon switching, they are more likely to return to a retailer they know of. Second, the 'Big 5' generally offer more expensive electricity tariffs than the other retailers. Therefore, consumers who have experienced cheaper alternatives in the past are more tempted to leave real-time pricing when there is a crisis because they are aware of them. While electricity is a homogeneous product, familiarity bias suggests that retailers can differentiate themselves. Regarding traditional tariffs, familiarity bias is consistent with the existence of search costs; consumers who don't search don't know alternative tariffs.

3.4.4 Discussion

We find that consumers who have more experience with real-time pricing are less likely to switch to another tariff during the crisis. The results suggest that real-time pricing is an experience good in the sense that it's a complex product consumers need to learn about to know whether it is indeed right for them. It is important to distinguish these experience effects from inertia effects were consumers "fall asleep" after switching. We do not think inertia effects play a role here for several reasons: The first is that, while more experienced consumers are indeed less likely to leave real-time pricing during the Winter 2017 crisis, more than 30% of those who do leave came back right after the crisis ended. This phenomenon cannot be explained by consumer inertia. Another reason why we do not think consumers become inattentive is that on RTP, they pay weekly bills and are regularly updated about any price spikes. This constant influx of information about market movements makes it unlikely that consumers are asleep. Finally, the Winter 2017 crisis was covered in the media quite extensively and therefore it seems unlikely that consumers were unaware it was happening. In conclusion, we argue that the decision to stay or switch was likely a conscious one.

The results also suggest that tariff complexity may be an issue: Consumers who had more experience with the tariff were less likely to leave and more likely to "game the system" by strategically leaving during the crisis before coming back afterwards. Another point supporting the suggestion that tariff complexity is important is the fact that spot prices were low at an unusually low volatility in the first 3 years after real-time pricing became available. Low and un-volatile spot prices are an ideal setting for consumers to earn about and experiment with real time pricing as potential losses are limited. Consumers joining right before Winter 2017 were not able to profit from this environment and may have been overwhelmed by the high spot prices, hence choosing to cut consumption or leave real-time pricing altogether. This factor suggests that the timing of real-time pricing roll-out may be important for consumers to stay on. Ideally, a policy maker or firm planning to introduce real-time pricing would like to do so in a period of low price- and volatility levels in order to make it possible for consumers to "ease into" the tariff. Interestingly, the retailer offering RTP plans, Flick Electric, started to guarantee new consumers that they would not pay more than what they would have paid under their previous retailer during the first year on real-time pricing. In essence, this policy offers consumers an insurance against large losses they may incur by experimenting with real-time pricing.

One reason that may have made consumers stay on real-time pricing during the Winter 2017 crisis is related to the regular updates they received from Flick Electric. In particular, consumers were able regularly updated on their cumulative savings since having joined Flick. These savings were computed by comparing their bills with a hypothetical bill had they stayed with their previous retailer[18]. Figure 3.4.5 provides two examples of how the information was displayed to consumers during Winter 2017 crisis. The consumer from the left panel had adopted RTP several months before the crisis and by beginning of June 2017 had accumulated more than $1500 in savings. The consumer from the right panel had adopted RTP just as the crisis started and after three weeks her savings had always been negative.

Unfortunately, our monthly consumption data does not allow us to calculate electricity bills in a reliable manner for consumers on real-time pricing. We have tried to build heuristic indicators

[18]For consumers who switched over from other providers offering non-standard contracts, the savings were calculated by comparing their bills to a hypothetical bill they would have faced by being with the retailer with the highest market share. For those consumers, the savings they were presented with in the app were probably overstated.

of savings by making ad-hoc assumptions on intra-day and intra-month consumption patterns and included them as an additional variable. The results are in Appendix H.2. However, given that we only have monthly consumption, we cannot account for intra-day and intra-month heterogeneity in consumer consumption patterns. Instead of directly relying on consumption, we have also regressed the probability to return to RTP on experience and a dummy equal to one if the consumer was originally with one of the five biggest retailers. The tariffs offered by those retailers tend to be more expensive, meaning in turn that consumers switching from such a tariff to RTP are more likely to run up large savings. However, we find no evidence that having been with a Big-5 retailer affects the decision to return to Flick after the crisis. As a consequence, our results shed little light on whether larger accumulated savings can help explain the decision to stay with real-time pricing during the Winter 2017 crisis.

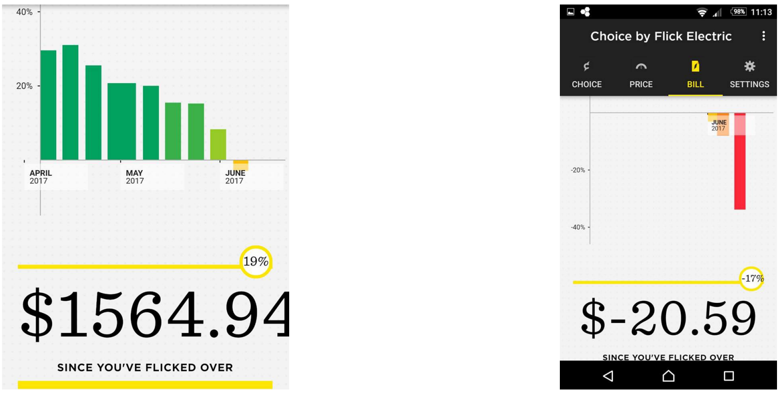


Figure 3.4.5: Screenshots (obtained online) of the display of two customers' cumulative savings on their mobile application.

3.5 Policy implications

Our results have several implications for policymakers. First, they offer an explanation why unraveling leading to large-scale adoption of real-time pricing did not occur in New Zealand. Second, the results can inform the debate on which measures to take in order to increase take-up of real-time pricing.

One key condition for unraveling to occur is that consumers need to self-identify as structural winners. They need to exhibit some form of forward-looking behavior that leads them conclude they will save by being on RTP without having to adjust their consumption profile. To the contrary, we find that consumers do not time their switch strategically but rather take one-shot decisions strongly influenced by contemporaneous spot prices. Simple heuristics such as current spot prices are unlikely to help in self-identifying as a structural winner. The adopters we observe are early adopters that differ strongly from the general population. Given that it is reasonable to assume that these early adopters are more interested in RTP and more sophisticated, later adopters would be even less likely to have the ability of self-identifying as structural winners.

Another condition for unraveling to occur is that once adopting, consumers stay on real-time pricing, even in periods when spot prices are high and volatile. This was not the case in New Zealand. While more experienced consumers were less likely to leave during the Winter 2017 crisis, less experienced consumers left in great numbers. What is more, those inexperienced consumers who left mostly did not return once spot prices had fallen. These findings suggest that inexperienced consumers who get a bad first impression upon adoption of RTP are likely to get scared and forego adopting RTP altogether in the future. Large numbers of consumers leaving RTP in periods of high spot prices make unraveling unlikely to occur.

A difficulty specific to opt-in policies is to convince consumers to adopt the tariff. Ito, Ida, and Tanaka (2017) show that providing information ex-ante to consumers significantly affects their tariff choices by helping the structural winners to self-select to time-varying tariffs. They also document self-selection on price-elasticity, suggesting that some consumers are aware of theirs ex ante. A simple policy to implement is to make it possible for consumers with smart meters to upload their real-time consumption in tariff comparison websites which exist in many countries already. Our finding that consumers use simple heuristics in their adoption decision suggests that such a provision of information may be helpful in increasing take-up of real-time pricing. Also, if consumers considering adopting RTP are excessively sensitive to contemporaneous events then one should design communication to steer them away. When spot prices are high, insist on long term benefits of adopting RTP or remind consumers about price seasonality. When spot prices are low, don't insist too much about the fact that they will increase at some point.

As discussed by Fowlie et al. (2020), if there is high inertia then policy-makers may consider defaulting some consumers to real-time pricing. This "opt-out" policy was adopted in Spain in 2014 but some evidence suggests that many consumers are not even aware of it. Our results show that bad first impressions can lead consumers who self-selected into the tariff to discard it eventually. This suggests that a major risk with opt-out policies is that, if it is poorly timed, all consumers get a bad first impression. This would surely make arguments in favor of real-time pricing politically untenable - and for a long time.

3.6 Conclusion

In this paper we document the adoption of a new electricity tariff, real-time pricing, by residential consumers in New Zealand. Contrary to theoretical predictions unraveling did not occur and, after more than seven years, less than 1.25% of consumers switched to this tariff. We find that prospective and recent adopters are highly sensitive to contemporaneous spot prices. The fact that inexperienced consumers focus on immediate outcomes combined with the fact that spot prices became volatile in winter 2017 and remained so afterwards could explain why few consumers adopted or remained on real-time pricing plans. However, after adopting the tariff, consumers become less responsive to ongoing events with experience. This finding calls for policies targeted at facilitating the learning process, before and after adoption, with the aim to steer attention away from immediate outcomes and towards long-term payoffs instead. First, consumers must be familiar with the formation of spot prices; the determinants of demand and supply, such as the weather, and in particular their seasonal patterns. Second, consumers need to be aware of whether, in the long run, they would benefit or not from switching to real-time pricing. Those who would benefit the most from it would switch first and the unraveling process would then follow. For that purpose, a sound policy would be to facilitate the access to records of household consumption profiles to use them on tariff comparison websites. This would help the structural winners - consumers who benefit from switching even without adjusting their consumption - to identify themselves. The other consumers would need to estimate the costs and benefits from adjusting their consumption. While they may be aware, partially or perfectly, of their price-elasticity they may need to learn about investments in energy efficient appliances. Finally, the overreaction to ongoing spot prices may also be explained by liquidity constraints because some consumers may not be able to pay large unexpected bills. While it seems unlikely that this issue was a major driver during the winter 2017 crisis in New Zealand, it may become one once the unraveling process progresses. This calls for insurance policies or financial instruments to help consumers smooth payments.

This paper opens several promising alleys for future research. If there are agency costs such as risk aversion (eg. liquidity constraints) or externalities (eg. social learning) then adoption rates will be inefficiency low. Too few consumers will willingly switch to real-time pricing and it may be necessary to encourage adoption and experimentation. Then, for instance, consumers could receive a transfer after they remained on real-time pricing a certain amount of time; and the transfer should not depend on the consumption while on the tariff, otherwise this would create distortions.[19] A retailer incurring the costs of experimentation may not be able to recover it ex post if, once the consumer has learned her valuation, she can switch to a competitor offering a low price. Thus, public subsidies may be required.

At least for as long as the unraveling process goes, and perhaps forever, some consumers will prefer tariffs with flat and fixed rates over a long horizon. In order to enforce long-term contracts, retailers usually set termination fees to discourage consumers from switching to

[19] In the case of New Zealand, the retailer Flick Electric guarantees its new customers that after 12-month they will have positive savings. This creates perverse incentives to over-consume during a crisis on the spot market.

another retailer before the end of the contract term. While termination fees protect the retailer from consumers arbitraging between their long-term tariffs when spot prices are high and other tariffs - such as real-time pricing - when spot prices are low, they also render consumers time-inelastic. Because inexperienced consumers overreact to contemporaneous spot prices, leaving them only short time-windows for adopting real-time pricing may interfere with the unraveling process. It may thus be necessary to reconsider how to enforce long-term contracts. In particular, because termination fees are often fixed and independent of how much time remains before the contract ends, which may not be efficient.

References

Hunt Allcott. Rethinking real-time electricity pricing. *Resource and energy economics*, 33(4): 820–842, 2011.

Soren T Anderson, Ryan Kellogg, and James M Sallee. What do consumers believe about future gasoline prices? *Journal of Environmental Economics and Management*, 66(3):383–403, 2013.

Severin Borenstein. The long-run efficiency of real-time electricity pricing. *The Energy Journal*, 26(3), 2005a.

Severin Borenstein. Time-varying retail electricity prices: Theory and practice. *Electricity deregulation: choices and challenges*, 4:317–356, 2005b.

Meghan R Busse, Devin G Pope, Jaren C Pope, and Jorge Silva-Risso. The psychological effect of weather on car purchases. *The Quarterly Journal of Economics*, 130(1):371–414, 2015.

Claude Crampes and Catherine Waddams. Empowering electricity consumers in retail and wholesale markets. *CERRE Report*, 2017.

Luisa Dressler and S Weiergraber. Alert the inert! switching costs and limited awareness in retail electricity markets. *Working Paper*, 2019.

Electricity Authority. 2017 Winter Review - Final report. 2018.

Natalia Fabra, David Rapson, Mar Reguant, and Jingyuan Wang. Estimating the elasticity to real time pricing: Evidence from the spanish electricity market. *American Economic Association Papers and Proceedings forthcoming*, 2021.

Meredith Fowlie, Catherine Wolfram, C Anna Spurlock, Annika Todd, Patrick Baylis, and Peter Cappers. Default effects and follow-on behavior: evidence from an electricity pricing program. NBER Working Papers 23553, National Bureau of Economic Research, 2020.

Michael D Grubb. Consumer inattention and bill-shock regulation. *The Review of Economic Studies*, 82(1):219–257, 2015.

Michael D Grubb and Matthew Osborne. Cellular service demand: Biased beliefs, learning, and bill shock. *American Economic Review*, 105(1):234–71, 2015.

David Hirshleifer, Ben Lourie, Thomas G Ruchti, and Phong Truong. First impression bias: Evidence from analyst forecasts. *Review of Finance*, 2020.

Ali Hortaçsu, Seyed Ali Madanizadeh, and Steven L Puller. Power to choose? an analysis of consumer inertia in the residential electricity market. *American Economic Journal: Economic Policy*, 9(4):192–226, 2017.

Koichiro Ito. Do consumers respond to marginal or average price? evidence from nonlinear electricity pricing. *American Economic Review*, 104(1):537–63, 2014.

Koichiro Ito, Takanori Ida, and Makoto Tanaka. Information Frictions, Inertia, and Selection on Elasticity: A Field Experiment on Electricity Tariff Choice. 2017.

Koichiro Ito, Takanori Ida, and Makoto Tanaka. Selection on welfare gains: Experimental evidence from electricity plan choice. NBER Working Papers 28413, National Bureau of Economic Research, 2021.

Paul L Joskow and Catherine D Wolfram. Dynamic pricing of electricity. *American Economic Review*, 102(3):381–85, 2012.

Stefan Lamp. Sunspots that matter. Working Paper, 2018.

Ulrike Malmendier and Stefan Nagel. Learning from inflation experiences. *The Quarterly Journal of Economics*, 131(1):53–87, 2016.

Ulrike Malmendier and Leslie Sheng Shen. Scarred consumption. NBER Working Papers 24696, National Bureau of Economic Research, 2019.

Eugenio J Miravete. Choosing the wrong calling plan? ignorance and learning. *American Economic Review*, 93(1):297–310, 2003.

Quang H Vuong. Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica: Journal of the Econometric Society*, pages 307–333, 1989.

Appendices

H.1 Regression tables

Table H.1.1: RTP-adoption, one-week period, data from 2014-06-01 to 2018-06-01

	Dependent variable:				
	Adopt RTP				
	(1)	(2)	(3)	(4)	(5)
Past Price	−0.077***			−0.112***	−0.066***
	(0.010)			(0.008)	(0.010)
Recent Price		−0.053***		−0.047***	−0.023**
		(0.009)		(0.009)	(0.011)
Future Price			−0.029***	−0.036***	−0.014
			(0.009)	(0.007)	(0.010)
Electricity consumption	0.037***	0.037***	0.037***		0.037***
	(0.003)	(0.003)	(0.003)		(0.003)
Seasonal consumption difference	−0.005	−0.005	−0.005		−0.005
	(0.033)	(0.033)	(0.033)		(0.033)
Income	1.601***	1.618***	1.619***		1.600***
	(0.463)	(0.463)	(0.463)		(0.463)
Age	−0.009***	−0.009***	−0.009***		−0.009***
	(0.001)	(0.001)	(0.001)		(0.001)
Work	0.551***	0.552***	0.550***		0.552***
	(0.114)	(0.113)	(0.113)		(0.114)
Education	1.259***	1.255***	1.256***		1.258***
	(0.122)	(0.122)	(0.122)		(0.122)
NRR FE?	Yes	Yes	Yes	No	Yes
Year-Month FE?	Yes	Yes	Yes	No	Yes
Retailer Origin FE?	Yes	Yes	Yes	No	Yes
Observations	143,435	143,435	143,435	149,479	143,435
Note:					*p<0.1; **p<0.05; ***p<0.01

Table H.1.2: RTP-adoption, four-week period, data from 2014-06-01 to 2018-06-01

	Dependent variable:				
	Adopt RTP				
	(1)	(2)	(3)	(4)	(5)
Past Price	-0.073^{***}			-0.059^{***}	-0.044^{***}
	(0.015)			(0.007)	(0.015)
Recent Price		-0.133^{***}		-0.187^{***}	-0.154^{***}
		(0.015)		(0.010)	(0.016)
Future Price			0.012	0.004	0.069^{***}
			(0.014)	(0.006)	(0.015)
Electricity consumption	0.037^{***}	0.037^{***}	0.037^{***}		0.037^{***}
	(0.003)	(0.003)	(0.003)		(0.003)
Seasonal consumption difference	-0.005	-0.004	-0.005		-0.004
	(0.033)	(0.033)	(0.033)		(0.033)
Income	1.607^{***}	1.587^{***}	1.625^{***}		1.580^{***}
	(0.463)	(0.463)	(0.463)		(0.463)
Age	-0.009^{***}	-0.009^{***}	-0.009^{***}		-0.009^{***}
	(0.001)	(0.001)	(0.001)		(0.001)
Work	0.549^{***}	0.552^{***}	0.550^{***}		0.553^{***}
	(0.113)	(0.114)	(0.113)		(0.114)
Education	1.258^{***}	1.257^{***}	1.256^{***}		1.259^{***}
	(0.122)	(0.122)	(0.122)		(0.122)
NRR FE?	Yes	Yes	Yes	No	Yes
Year-Month FE?	Yes	Yes	Yes	No	Yes
Retailer Origin FE?	Yes	Yes	Yes	No	Yes
Observations	143,435	143,435	143,435	149,479	143,435
Note:					*p<0.1; **p<0.05; ***p<0.01

Table H.1.3: RTP-adoption - alternative price definitions - data from 2014-06-01 to 2018-06-01

	Dependent variable:					
	Adopt RTP					
	1 Week	2 Weeks	4 Weeks	1 Week	2 Weeks	4 Weeks
	(1)	(2)	(3)	(4)	(5)	(6)
Recent Price	−0.163***	−0.191***	−0.229***	−0.063***	−0.098***	−0.145***
	(0.005)	(0.005)	(0.006)	(0.009)	(0.011)	(0.014)
Last Year Price	−0.0001	0.007	0.021***	0.014	0.024	0.002
	(0.007)	(0.007)	(0.008)	(0.015)	(0.018)	(0.024)
Future Predicted Price	−0.004	−0.004	−0.0004	−0.028***	−0.031***	−0.040***
	(0.003)	(0.003)	(0.004)	(0.005)	(0.005)	(0.006)
Electricity consumption				0.032***	0.032***	0.032***
				(0.003)	(0.003)	(0.003)
Seasonal consumption difference				0.130***	0.130***	0.131***
				(0.032)	(0.032)	(0.032)
Income				0.002***	0.002***	0.002***
				(0.0005)	(0.0005)	(0.0005)
Age				−0.010***	−0.010***	−0.010***
				(0.001)	(0.001)	(0.001)
Work				0.480***	0.480***	0.480***
				(0.111)	(0.112)	(0.112)
Education				1.246***	1.249***	1.251***
				(0.119)	(0.119)	(0.119)
NRR FE?	No	No	No	Yes	Yes	Yes
Year-Month FE?	No	No	No	Yes	Yes	Yes
Retailer Origin FE?	No	No	No	Yes	Yes	Yes
Observations	149,479	149,479	149,479	149,479	149,479	149,479
Note:						*p<0.1; **p<0.05; ***p<0.01

Table H.1.4: Discarding RTP during Winter 2017 crisis: robustness check

	Dependent variable:				
	Discard RTP				
	(1)	(2)	(3)	(4)	(5)
Time on RTP (month)	−0.03***	−0.02***	−0.02***	−0.02***	−0.03***
	(0.005)	(0.01)	(0.01)	(0.01)	(0.01)
Joined Last	0.41***	0.46***	0.36***	0.23***	0.01
	(0.10)	(0.08)	(0.08)	(0.09)	(0.10)
Annual elec cons (MWh)	−0.02**	−0.02**	−0.02**	−0.02**	−0.02**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Win/Sum cons diff (MWh)	0.39***	0.39***	0.40***	0.40***	0.39***
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Income (k$/yr)	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	−0.01**	−0.01**	−0.01**	−0.01**	−0.01**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Work (%)	−0.005	−0.01	−0.005	−0.005	−0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Education (%)	−0.004	−0.004	−0.004	−0.003	−0.003
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Last Cohort	1 month	2 months	3 months	6 months	9 months
Location-on-Previous retailer FE?	Yes	Yes	Yes	Yes	Yes
Observations	8,617	8,617	8,617	8,617	8,617
Log Likelihood	−4,013.64	−4,007.64	−4,012.67	−4,019.36	−4,022.85
Akaike Inf. Crit.	8,127.28	8,115.29	8,125.34	8,138.71	8,145.70
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table H.1.5: Probability of discarding RTP

	Dependent variable:					
	Discard RTP					
	Wel-Chr	Wel-Chr	Wel-Chr	Auc-Chr	Auc-Chr	Auc-Chr
	(1)	(2)	(3)	(4)	(5)	(6)
Time on RTP (month)	-0.04^{***}	-0.04^{***}	-0.04^{***}	-0.05^{**}	-0.04^{*}	-0.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
Christchurch	0.18	0.15	0.30	0.13	0.10	0.46
	(0.14)	(0.14)	(0.34)	(0.20)	(0.20)	(0.55)
Time on RTP x Christchurch	-0.004	0.003	0.01	-0.001	0.005	-0.01
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Annual elec cons (MWh)	-0.02^{**}	-0.02^{**}	-0.02^{**}	-0.02^{**}	-0.02^{**}	-0.02^{**}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Win/Sum cons diff (MWh)	0.41^{***}	0.40^{***}	0.40^{***}	0.46^{***}	0.45^{***}	0.45^{***}
	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)
Income (k$/yr)	0.001	0.001	0.001	-0.001	-0.0004	-0.0004
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
Age	-0.01^{**}	-0.01^{**}	-0.01^{**}	-0.01	-0.01	-0.01
	(0.005)	(0.005)	(0.005)	(0.01)	(0.01)	(0.01)
Work (%)	-0.01^{*}	-0.01^{*}	-0.01^{*}	-0.004	-0.004	-0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Education (%)	-0.002	-0.002	-0.002	-0.005	-0.01	-0.01
	(0.004)	(0.004)	(0.004)	(0.01)	(0.01)	(0.01)
Month FE?	No	Yes	No	No	Yes	No
Month-on-NRR FE?	No	No	Yes	No	No	Yes
Observations	6,142	6,142	6,142	4,923	4,923	4,923
Log Likelihood	$-3,109.52$	$-3,098.54$	$-3,094.68$	$-2,575.69$	$-2,563.30$	$-2,558.82$
Akaike Inf. Crit.	6,239.04	6,239.09	6,253.36	5,171.37	5,168.60	5,181.64
Note:						*p<0.1; **p<0.05; ***p<0.01

Table H.1.6: Switching back to RTP after the Winter 2017 crisis as a function of experience

	Dependent variable:		
	Return to RTP		
	(1)	(2)	(3)
Time on RTP (month)	0.07***	0.07***	0.08***
	(0.01)	(0.01)	(0.01)
Annual elec cons (MWh)	−0.003	−0.002	−0.004
	(0.01)	(0.01)	(0.01)
Win/Sum cons diff (MWh)	0.03	0.0003	−0.01
	(0.13)	(0.13)	(0.13)
Income (k$/yr)	0.004*	0.004*	0.004
	(0.002)	(0.002)	(0.002)
Age	0.002	0.003	0.004
	(0.01)	(0.01)	(0.01)
Work (%)	0.002	0.002	0.002
	(0.01)	(0.01)	(0.01)
Education (%)	0.001	0.0004	−0.00004
	(0.01)	(0.01)	(0.01)
Location FE?	Yes	Yes	No
Previous retailer FE?	No	Yes	No
Location-on-Previous retailer FE?	No	No	Yes
Observations	2,339	2,339	2,339
Log Likelihood	−1,288.16	−1,275.54	−1,262.47
Akaike Inf. Crit.	2,596.33	2,603.08	2,618.93
Note:	*p<0.1; **p<0.05; ***p<0.01		

Table H.1.7: Switching back to RTP after the Winter 2017 crisis as a function of experience - TEST

	Dependent variable:
	Return to RTP
Time on RTP (month)	0.07***
	(0.01)
Origin Big 5	−0.02
	(0.11)
Annual elec cons (MWh)	−0.003
	(0.01)
Win/Sum cons diff (MWh)	0.03
	(0.13)
Income (k$/yr)	0.004*
	(0.002)
Age	0.002
	(0.01)
Work (%)	0.002
	(0.01)
Education (%)	0.001
	(0.01)
Location FE?	Yes
Observations	2,339
Log Likelihood	−1,288.14
Akaike Inf. Crit.	2,598.28
Note:	*p<0.1; **p<0.05; ***p<0.01

	Average	Big 5	Non-Traditional	Others
Origin of RTP-adopters (%)		73.6	17.5	8.8
Switch during Winter 2017 crisis (%)	13.6	12.5	17.6	15.6
Switch to Big 5 (%)	39.8	46.8	24.4	27.7
Switch to Non-Traditional (%)	47.4	42.6	63.9	42.3
Switch to Others (%)	12.8	10.5	11.7	30.0

Table H.1.11: Previous experience affects switching decisions - Retailer of origin

Table H.1.8: Comparison of consumption reduction, consumers on RTP vs other consumers

	Dependent variable:					
	Consumption					
	2014	2014	2015	2015	Mean 14 + 15	Mean 14 + 15
	(1)	(2)	(3)	(4)	(5)	(6)
Post	2.746***	2.746***	−27.420***	−27.420***	−12.337***	−12.337***
	(0.974)	(0.927)	(0.994)	(0.947)	(0.958)	(0.910)
RTP	131.251***	71.421***	139.495***	78.701***	135.373***	75.061***
	(5.135)	(4.857)	(5.359)	(5.054)	(4.909)	(4.595)
I(Post *RTP)	−11.404	−11.404*	−19.649***	−19.649***	−15.526**	−15.526**
	(7.184)	(6.812)	(7.346)	(6.955)	(7.025)	(6.629)
MedianRevenue		0.004***		0.004***		0.004***
		(0.00002)		(0.00002)		(0.00002)
MedianAge		0.875***		0.782***		0.828***
		(0.062)		(0.063)		(0.061)
WorkCategoryManPro		−0.057		−0.049		−0.053
		(0.039)		(0.040)		(0.038)
EducHigh		−2.933***		−2.991***		−2.962***
		(0.051)		(0.052)		(0.050)
factor(NRR)23		58.292***		56.527***		57.409***
		(1.101)		(1.123)		(1.083)
factor(NRR)30		294.352***		302.348***		298.350***
		(1.257)		(1.284)		(1.228)
Observations	1,013,944	1,013,944	1,013,944	1,013,944	1,013,944	1,013,944
R^2	0.001	0.094	0.002	0.094	0.002	0.099
Adjusted R^2	0.001	0.094	0.002	0.094	0.002	0.099
Residual Std. Error	485.681	462.626	495.777	472.283	477.822	453.888
F Statistic	455.840***	11,676.770***	732.138***	11,757.250***	544.187***	12,396.550***
Note:						*p<0.1; **p<0.05; ***p<0.01

Table H.1.9: Consumption reduction as a function of experience

	Dependent variable:					
	ConsW17	log(ConsW17)	ConsW17	log(ConsW17)	ConsW17	log(ConsW17)
	(1)	(2)	(3)	(4)	(5)	(6)
I(Experience/30)	3.135***	0.004***	2.849***	0.004***	2.696***	0.003***
	(0.543)	(0.001)	(0.501)	(0.001)	(0.493)	(0.001)
ConsW14	0.546***				0.241***	
	(0.011)				(0.015)	
log(ConsW14)		0.443***				0.247***
		(0.013)				(0.016)
ConsW15			0.575***		0.410***	
			(0.011)		(0.015)	
log(ConsW15)				0.467***		0.328***
				(0.016)		(0.018)
MedianRevenue	0.002***	0.00000***	0.002***	0.00000***	0.001***	0.00000***
	(0.0002)	(0.00000)	(0.0002)	(0.00000)	(0.0002)	(0.00000)
MedianAge	1.078*	0.001	1.050*	0.001	0.842	0.001
	(0.581)	(0.001)	(0.548)	(0.001)	(0.534)	(0.001)
WorkCategoryManPro	1.368***	0.002***	1.108***	0.002***	1.125***	0.002***
	(0.438)	(0.001)	(0.383)	(0.0005)	(0.379)	(0.0005)
EducHigh	−2.610***	−0.004***	−2.315***	−0.004***	−2.114***	−0.003***
	(0.456)	(0.001)	(0.419)	(0.001)	(0.412)	(0.001)
factor(NRR)23	55.453***	0.084***	58.887***	0.088***	48.985***	0.069***
	(9.867)	(0.013)	(9.431)	(0.013)	(9.230)	(0.012)
factor(NRR)30	133.283***	0.193***	105.899***	0.181***	81.024***	0.133***
	(12.228)	(0.016)	(11.746)	(0.016)	(11.501)	(0.015)
Observations	10,373	10,373	10,373	10,373	10,373	10,373
R^2	0.376	0.303	0.431	0.339	0.457	0.381
Adjusted R^2	0.375	0.302	0.431	0.339	0.457	0.381
Residual Std. Error	400.712	0.481	382.442	0.468	373.591	0.453
F Statistic	778.969***	562.681***	981.898***	664.596***	969.971***	709.910***
Note:						*p<0.1; **p<0.05; ***p<0.01

Table H.1.10: Consumption reduction as a function of experience and origin retailer

	Dependent variable:					
	ConsW17	log(ConsW17)	ConsW17	log(ConsW17)	ConsW17	log(ConsW17)
	(1)	(2)	(3)	(4)	(5)	(6)
I(Experience/30)	3.107***	0.004***	2.827***	0.004***	2.674***	0.003***
	(0.543)	(0.001)	(0.501)	(0.001)	(0.494)	(0.001)
Big5	−16.023*	−0.018	−12.327	−0.013	−12.552	−0.014
	(9.392)	(0.011)	(8.752)	(0.011)	(8.609)	(0.010)
ConsW14	0.546***				0.241***	
	(0.011)				(0.015)	
log(ConsW14)		0.443***				0.247***
		(0.013)				(0.016)
ConsW15			0.575***		0.410***	
			(0.011)		(0.015)	
log(ConsW15)				0.467***		0.328***
				(0.016)		(0.018)
MedianRevenue	0.002***	0.00000***	0.002***	0.00000***	0.001***	0.00000***
	(0.0002)	(0.00000)	(0.0002)	(0.00000)	(0.0002)	(0.00000)
MedianAge	1.098*	0.001	1.064*	0.001	0.857	0.001
	(0.581)	(0.001)	(0.547)	(0.001)	(0.534)	(0.001)
WorkCategoryManPro	1.372***	0.002***	1.111***	0.002***	1.128***	0.002***
	(0.438)	(0.001)	(0.383)	(0.0005)	(0.379)	(0.0005)
EducHigh	−2.606***	−0.004***	−2.312***	−0.004***	−2.111***	−0.003***
	(0.456)	(0.001)	(0.419)	(0.001)	(0.412)	(0.001)
factor(NRR)23	58.993***	0.088***	61.614***	0.091***	51.761***	0.073***
	(10.061)	(0.013)	(9.593)	(0.013)	(9.391)	(0.012)
factor(NRR)30	137.393***	0.198***	109.081***	0.185***	84.261***	0.136***
	(12.566)	(0.016)	(12.012)	(0.016)	(11.785)	(0.015)
Observations	10,373	10,373	10,373	10,373	10,373	10,373
R^2	0.376	0.303	0.431	0.339	0.457	0.381
Adjusted R^2	0.375	0.302	0.431	0.339	0.457	0.381
Residual Std. Error	400.677	0.480	382.428	0.468	373.574	0.453
F Statistic	692.844***	500.517***	873.066***	590.924***	873.251***	639.113***
Note:						*p<0.1; **p<0.05; ***p<0.01

H.2 Discarding RTP and savings

Table H.2.1: Probability of discarding RTP: the role of cumulated savings

	Dependent variable:	
	Discard RTP	
	Uniform	Peak
	(1)	(2)
Past Savings	-0.0001	0.008^{***}
	(0.0002)	(0.001)
Time on RTP (months)	-0.001	-0.010^{***}
	(0.001)	(0.001)
Income	-0.001^{***}	-0.003^{***}
	(0.0004)	(0.0003)
Age	0.001	-0.001
	(0.002)	(0.002)
Work	-0.024^{***}	-0.023^{***}
	(0.006)	(0.007)
Education	-0.004	-0.004
	(0.003)	(0.003)
Wellington	-0.003	-0.004
	(0.005)	(0.005)
Christchurch	0.081	-0.696^{***}
	(0.161)	(0.172)
factor(NRR)30	0.316^{*}	-0.992^{***}
	(0.178)	(0.177)
Observations	3,111	3,111
Log Likelihood	$-1,481.141$	$-1,329.711$
Akaike Inf. Crit.	2,982.282	2,679.423
Note:	*p<0.1; **p<0.05; ***p<0.01	

H.3 Short regressions

Table H.3.1: Discarding RTP during Winter 2017 crisis: selection vs experience

	Dependent variable:					
	Discard RTP					
	(1)	(2)	(3)	(4)	(5)	(6)
Time on RTP (months)	−0.03***		−0.03***		−0.04***	
	(0.004)		(0.005)		(0.005)	
Income	0.0005	0.0004	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age	−0.01***	−0.01***	−0.01**	−0.01**	−0.01**	−0.01**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Work status	−0.01	−0.01*	−0.01	−0.01*	−0.005	−0.01
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Education	−0.003	−0.003	−0.003	−0.004	−0.003	−0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Location FE?	Yes	Yes	Yes	Yes	No	No
Previous retailer FE?	No	No	Yes	Yes	No	No
Location-on-Previous retailer FE?	No	No	No	Yes	Yes	
Observations	8,617	8,617	8,617	8,617	8,617	8,617
Log Likelihood	−4,085.61	−4,116.60	−4,039.31	−4,068.43	−4,022.86	−4,052.50
Akaike Inf. Crit.	8,191.21	8,251.20	8,132.62	8,188.86	8,143.72	8,201.00
Note:						*p<0.1; **p<0.05; ***p<0.01

Table H.3.2: Discarding RTP during Winter 2017 crisis: robustness check

	Dependent variable:				
	Discard RTP				
	(1)	(2)	(3)	(4)	(5)
Time on RTP (month)	−0.03***	−0.02***	−0.02***	−0.02***	−0.03***
	(0.005)	(0.01)	(0.01)	(0.01)	(0.01)
Joined Last	0.41***	0.46***	0.36***	0.23***	0.01
	(0.10)	(0.08)	(0.08)	(0.09)	(0.10)
Last Cohort	1 month	2 months	3 months	6 months	9 months
Location-on-Previous retailer FE?	Yes	Yes	Yes	Yes	Yes
Observations	8,617	8,617	8,617	8,617	8,617
Log Likelihood	−4,013.64	−4,007.64	−4,012.67	−4,019.36	−4,022.85
Akaike Inf. Crit.	8,127.28	8,115.29	8,125.34	8,138.71	8,145.70
Note:					*p<0.1; **p<0.05; ***p<0.01

Table H.3.3: Probability of discarding RTP: Wellington vs Christchurch, only observations after RTP is available in Christchurch

	Dependent variable:		
	Discard RTP		
	(1)	(2)	(3)
Time on RTP (months)	−0.04***	−0.04***	−0.04***
	(0.01)	(0.01)	(0.01)
Christchurch	0.18	0.15	0.30
	(0.14)	(0.14)	(0.34)
Time on RTP x Christchurch	−0.004	0.003	0.01
	(0.01)	(0.02)	(0.02)
Month FE?	No	Yes	No
Month-on-NRR FE?	No	No	Yes
Observations	6,142	6,142	6,142
Log Likelihood	−3,109.52	−3,098.54	−3,094.68
Note:			*p<0.1; **p<0.05; ***p<0.01

Table H.3.4: Consumption reduction as a function of experience

	Dependent variable:					
	ConsW17	log(ConsW17)	ConsW17	log(ConsW17)	ConsW17	log(ConsW17)
	(1)	(2)	(3)	(4)	(5)	(6)
Time on RTP (months)	3.135***	0.004***	2.849***	0.004***	2.696***	0.003***
	(0.543)	(0.001)	(0.501)	(0.001)	(0.493)	(0.001)
Consumption 2014	0.546***				0.241***	
	(0.011)				(0.015)	
log(Consumption 2014)		0.443***				0.247***
		(0.013)				(0.016)
Consumption 2015			0.575***		0.410***	
			(0.011)		(0.015)	
log(Consumption 2015)				0.467***		0.328***
				(0.016)		(0.018)
Observations	10,373	10,373	10,373	10,373	10,373	10,373
R^2	0.376	0.303	0.431	0.339	0.457	0.381
Adjusted R^2	0.375	0.302	0.431	0.339	0.457	0.381
Residual Std. Error	400.712	0.481	382.442	0.468	373.591	0.453
F Statistic	778.969***	562.681***	981.898***	664.596***	969.971***	709.910***
Note:						*p<0.1; **p<0.05; ***p<0.01

Table H.3.5: RTP-adoption, two-week period, data from 2014-06-01 to 2018-06-01

	Dependent variable:				
	Adopt RTP				
	(1)	(2)	(3)	(4)	(5)
Past Price	-0.084^{***}			-0.139^{***}	-0.068^{***}
	(0.011)			(0.008)	(0.012)
Recent Price		-0.089^{***}		-0.083^{***}	-0.075^{***}
		(0.011)		(0.009)	(0.012)
Future Price			-0.020^{*}	-0.006	0.009
			(0.011)	(0.007)	(0.012)
I(Cons2014/1000)	0.037^{***}	0.037^{***}	0.037^{***}		0.037^{***}
	(0.003)	(0.003)	(0.003)		(0.003)
I(WSConsDiff14/1000)	-0.005	-0.004	-0.005		-0.004
	(0.033)	(0.033)	(0.033)		(0.033)
NRR FE?	Yes	Yes	Yes	No	Yes
Year-Month FE?	Yes	Yes	Yes	No	Yes
Retailer Origin FE?	Yes	Yes	Yes	No	Yes
Observations	143,435	143,435	143,435	149,479	143,435
Note:					*p<0.1; **p<0.05; ***p<0.01