THÈSE de DOCTORAT

de l'UNIVERSITÉ TOULOUSE CAPITOLE

Présentée et soutenue par

Madame Léa BIGNON

Le 10 juillet 2025

Essais en économie industrielle et économie de la santé

École doctorale : **TSE Toulouse Sciences Economiques**
Spécialité : **Sciences Economiques - Toulouse**
Unité de recherche : **TSE-R - Toulouse School of Economics - Recherche**

Thèse dirigée par *Monsieur Pierre DUBOIS*

Composition du jury
Rapporteure : Mme Amanda STARC
Rapporteure : Mme Ashley SWANSON
Examinateur : M. Bruno JULLIEN
Directeur de thèse : M. Pierre DUBOIS

Essays in Industrial Organization and Health Economics

Léa Bignon

May 21, 2025

Acknowledgement

First and foremost, I am infinitely grateful to my advisor, Pierre Dubois, for his unwavering support. His encouragement and confidence were instrumental in helping me carry out this thesis. I am eternally thankful for his time and thoughtful guidance.

I am also profoundly grateful to Bruno Jullien for his continuous support over the years. Bruno has always been generous with his time. I greatly benefited from his theoretical perspective and our many conversations.

I would like to thank Matthew Backus and Matthew Grennan. I have benefited tremendously from Matt Backus's insights, support, and invaluable advice throughout the Job Market process. I am grateful to Matt Grennan for inviting me to UC Berkeley Haas School of Business, for his thoughtful questions and comments, which helped me substantially shape my Job Market Paper. I would also like to thank Christian Hellwig for his time and precious feedback on my paper.

I thank Amanda Starc and Ashley Swanson for agreeing to referee my thesis.

I am grateful to my co-authors, Alessandro Iaria and Laura Lasio. This thesis would have never started without Alessandro's enthusiasm for teaching Empirical Industrial Organization.

Many people took the time to listen to my research throughout the years, from faculty members at TSE and UC Berkeley to seminar speakers. I would like to thank all of them, in particular Christian Bontemps, Paul Diegert, Isis Durrmeyer, Daniel Ershov, Ben Handel, Chuqing Jin, Thierry Magnac, Mathias Reynaert, François Salanié, and Carolyn Stein. Each of them contributed to this thesis, helping me grow as a researcher. I am also thankful to all members of the Industrial Organization and Econometrics groups at TSE and the Industrial Organization group at UC Berkeley for their feedback and comments.

I thank the Toulouse School of Economics for providing an intellectually stimulating and supportive environment and the UC Berkeley Haas School of Business for its hospitality. In particular, I thank Nour Meddahi and Fabrice Collard for their work as heads of the PhD program, Ana Gazmuri, Uli Hege, and Louise Strachan for their help during the Job Market process, and Marie-Hélène Dufour and Alexandro Noceto for their administrative and logistical support.

I would like to thank all the people I met throughout my PhD, from Toulouse to Berkeley – especially Alfonso, Amirreza, Amory, Anaïs, Chiara, David, Gilles, Giulia, Gökçe, Gosia, Guillem, Gyung-Mo, Hippolyte, J-F, Jeremy, Jeff, Johanna, Juan, Lilian, Louis, Lony, Luisa,

Luise, Lisa, Maria, Marion, Max, Moritz, Mudit, Paula, Pedram, Peter, Rossi, Simon, Stephan, Tanja, Tim, Tuuli, Valentina, Wenxuan, and Xin. Thank you for the laugh, the support, the dinners, drinks, coffees, basketball games, climbing, hiking, skiing, and (attempts) surfing. This adventure would not have been as enjoyable without you.

This journey started way before Toulouse and I would like to thank my friends from Brittany and Paris, and especially Alexis, Benoit, Bob, Etienne, Karen, Mathilde, Rémi, Richard, Théo, Amandine, Antonin G., Baptiste, and Julia for being constant reminders that life is much more than academic research.

Finally, I would like to thank my family. To my parents for always supporting me. To my sisters, Agathe and Juliette, for their constant encouragement and endless support.

Summary

This thesis studies Industrial Organization questions arising in pharmaceutical markets where *(i)* new technologies increase individual-level information available for decision making; *(ii)* private insurers' incentives can undermine consumer protection; and *(iii)* the cost structure of new treatments challenges traditional pricing rules.

In the first chapter, entitled *"Pharmaceuticals and Digital Health: How Data-driven Insights May Reshape the Insulin Market"*, I document cross-market complementarities between medical devices and pharmaceuticals – how innovative medical devices, generating individual-level health data, can shape demand, pricing, and innovation incentives in pharmaceutical markets. I study these questions by analyzing the impact of Continuous Glucose Monitors (CGMs) on the insulin market, where the high-frequency data from CGMs can reduce the information frictions that physicians face about drugs. Using French health insurance claims data, I show that CGMs shift drug choices through new observable attributes (e.g. overnight performance). The device allows the manufacturers of drugs with strong performance on newly observable attributes to benefit from higher demand and set higher prices. My findings indicate that the introduction of these new attributes into pharmaceutical demand can alter the relative profitability of drug innovation strategies, thereby shaping future pharmaceutical innovation.

In the second chapter, entitled *"Strategic Tier Design in Health Insurance: The Case of Medicare Part D"* and co-authored with Laura Lasio and Alessandro Iaria, we study how Medicare Part D insurers strategically manipulate drug tiers to increase generic drugs' cost sharing. We show that plans increasingly moved generics to higher tiers, raising out-of-pocket costs despite regulatory constraints. Exploiting institutional rules and regional variation, we estimate that this "tier upgrading" raised generic spending by up to $100 per enrollee and reduced utilization. We find that welfare losses include reduced adherence to high-value drugs.

In the third chapter, entitled *"Biologic Drugs and Learning-by-Doing"* and co-authored with Pierre Dubois, we measure the extent of learning-by-doing in the production process of biologic drugs, which are complex therapies. Using international drug sales data, we estimate a model of demand, pricing, and production for biologic drugs. We find that the marginal costs of biosimilars decreased by 46% since market entry, partly due to returns to experience. As biologic drugs production is geographically concentrated, learning-by-doing creates inter-temporal and international spillovers in cost efficiencies. Delayed biosimilar entry in one country raises costs in others. These dynamics challenge static drug pricing rules.

Contents

Chapter 1

Pharmaceuticals and Digital Health: How Data-driven Insights May Reshape the Insulin Market

Abstract

Digital health technologies, such as Continuous Glucose Monitors (CGMs), are transforming the availability of patient-level data, potentially influencing healthcare markets more broadly. This paper examines how CGMs influence the insulin market, shedding light on the impact of digital health technologies on pharmaceutical demand, pricing, and innovation incentives. I develop and estimate a tractable model of supply and demand for insulin, embedding: (i) patient-specific learning about treatment performance through the digital device, (ii) physician-level learning about new insulin products based on patient experiences, and (iii) price bargaining by pharmaceutical companies and the regulator, both internalizing demand-side learning. Using comprehensive medical claims data from France, where expanded CGM coverage boosted technology adoption, I find that CGMs' patient-specific information steered insulin demand towards newer products, with limited information spillover to nonusers. Manufacturers of drugs that benefited from higher perceived quality, thanks to the observability of these attributes, could negotiate higher drug prices. My findings indicate that the introduction of these new *observable* attributes into pharmaceutical demand can shift the relative profitability of drug innovation strategies, thereby shaping future pharmaceutical innovation.

1.1 Introduction

From smartphone step counters to smartwatches, digital devices that generate high-frequency health data are now widely available, transforming the landscape of individual-level information for decision-making.[1] These data are compelling for the pharmaceutical industry because prescription drugs are experience goods. Patients exhibit diverse needs, leading to idiosyncratic match values with drugs, and physicians learn about a drug's performance as they prescribe it. Therefore, by generating timely information, digital technologies can also affect the demand for other healthcare products.

This paper investigates the impact of Continuous Glucose Monitors (CGMs) on insulin choice in diabetes treatment. By delivering continuous glucose readings, CGMs enhance the information available to evaluate insulin treatments, highlighting the limitations of traditional measures that may conceal critical variations in glucose control (Figure 1.19). Key questions arise: How do CGMs' insights influence the insulin choices for technology users? To what extent do these insights guide physicians' prescriptions for other patients? What are the implications for pharmaceutical price negotiations between manufacturers and regulators? Finally, how do CGMs impact incentives for pharmaceutical innovation?

To address these questions, I leverage comprehensive medical claims data from the French health insurance system, which provides a unique setup to assess the impacts of CGMs. This dataset is particularly valuable due to France's universal, centralized insurance system, which provides comprehensive records of prescription reimbursements for the entire population. Additionally, a policy change that expanded CGM coverage boosted adoption among insulin-dependent diabetic patients. This policy shift, along with the technological characteristics of the device, allows for inferring technology adoption and attrition from claims data.[2] I use the claims data to estimate a tractable demand and supply model for insulin to show that patient-specific information from CGMs steers insulin demand towards newer, less familiar drugs, fostering physicians' learning about these products' real-life performance. By influencing how physicians perceive drugs' clinical match values, CGMs enable the manufacturers of drugs that perform well on the new *observable* attributes to set higher prices, ultimately impacting the

[1] Examples of such technologies include Continuous Glucose Monitors (CGMs) measuring sugar levels, wearable blood pressure trackers, connected insertable cardiac monitors, and sleep trackers (Handel and Kolstad (2017)).

[2] The adoption date for patients paying for CGMs out-of-pocket before the coverage decision can be inferred through claims for alternative glucose measurement technologies. The current technology relies on disposable sensors that must be replaced every 14 days.

profitability of the pharmaceutical innovation strategy.

This paper sheds light on how digital health technologies shape pharmaceutical markets, disentangling two mechanisms. First, high-frequency data can speed up physicians' learning about the performance of treatments. Second, CGMs provide detailed patient-specific insights, broadening the attributes observable for evaluating the effectiveness of drugs. This distinction matters: faster learning upon market entry dampens the barriers to new drug adoption without changing the perceived differentiation of drugs after the initial uncertainty resolves. In contrast, new attributes for evaluating drug performance change the information available to physicians when choosing treatments beyond market entry. By reshaping physicians' preferences for treatments, the attributes observed thanks to these new technologies may not only impact competition between existing products but also affect the incentives for future pharmaceutical innovations. I develop an empirical framework of supply and demand for insulin, accounting for: (i) patient-specific match value components revealed by CGMs, (ii) physician-level learning about new drugs, and (iii) price responses from bargaining with the regulator. Using claims data from 2015 to 2021 — a period marked by new insulin entries and increased CGM adoption — I identify (i) the impact of CGM-generated patient-specific information by comparing insulin choices of similar patients with and without CGMs and (ii) the impact of CGMs on physicians' learning speed through variations in prescription patterns as they gain experience with drugs from patients with and without CGMs.

In a descriptive analysis, I provide four empirical facts highlighting the interaction between CGMs and insulin demand. First, CGM adoption is widespread among diabetic patients using long-acting insulin — enough to significantly affect insulin demand. Approximately 34% of existing patients adopted CGM. Second, poor diabetes control *ex-ante* does not predict CGM adoption. Hospitalizations and ER visits before CGM coverage account for only 1.5% of the explained variation in technology adoption. Convenience is the main driver of CGM adoption as glucose control without CGM is burdensome. Third, CGM adoption impacts insulin treatment choice. CGM users are 5 to 7.7 percentage points more likely to switch treatment shortly after adoption compared to nonusers, a correlation that disappears in the medium or long run. Patients switched from older insulin products to newer alternatives, allowing physicians to gain more experience with these drugs. Fourth, physicians are learning about the real-life performance of new drugs across patients. As they see more patients using new drugs in real-life conditions, they are more likely to switch more patients to that treatment.

The insulin market changed, regardless of CGMs, due to new product entries before and after the CGM insurance coverage decision. A structural model is necessary to disentangle the effect of new product entries separately from CGM adoption and pinpoint the impact of patient-specific information from physician-level learning. Motivated by the empirical facts, I develop a tractable framework in which, on the demand side, physicians facing heterogeneous patients choose the most cost-effective insulin. For patients without CGMs, incomplete information about clinical match values — especially for new products — affects choices. Physician learning occurs through direct patient experience (Coscelli and Shum (2004)) but remains incomplete without patient-specific CGM data. The device generates unique insights about the patient's glucose profile, which can be informative about (i) the patient, independent of the medication; (ii) the drug, independent of the patient; or (iii) the patient-drug combination. However, only the latter two channels affect pharmaceutical demand, implying that some complementarity between CGM insights and the drug's mechanism of action is crucial in influencing demand. In my empirical model, insights from CGMs affect insulin demand from users and alter physicians' dynamic learning about new products' performance. On the supply side, the regulator and insulin manufacturers bargain over price, internalizing demand-side learning and perceived product differentiation. CGMs potentially influence equilibrium insulin prices by impacting new drug learning and perceived product differentiation.

I estimate the demand model via simulated maximum likelihood to account for unobserved patient experience. To address the selection of patients into CGM, I use a flexible specification for the patient-product match values, controlling for product-specific unobserved heterogeneity common across similar patients. The remaining within-group patient-product heterogeneity is assumed to be independent of the drivers of CGM adoption. The estimates suggest pessimistic beliefs about the performance of new products upon entry, similar to the findings of Coscelli and Shum (2004) for the anti-ulcer drugs market in Italy. CGM influences insulin choice through an idiosyncratic match value component observed for users thanks to the device. It does not increase or decrease the precision of the experience signal about new drugs that a physician receives during a medical appointment, suggesting limited information spillovers to nonusers. The supply model is estimated using GMM. As equilibrium prices remain relatively low despite quasi-inelastic insulin demand, most bargaining weight can be attributed to the regulator who maximizes consumer surplus.[3]

[3]In countries like the US, where the government did not intervene in drug pricing until recently, insulin prices for the same products as those offered in France are significantly higher.

Finally, I use the demand and supply framework estimates to compute two counterfactuals illustrating how digital devices affect pharmaceutical markets. First, I assess the short-term impact of CGMs on market outcomes by simulating the equilibrium prices and market shares in the absence of CGMs, holding the set of existing products constant. Second, I explore the potential long-term implications of CGMs on pharmaceutical innovation by simulating the entry of alternative drugs under different technological environments — both with and without CGMs.

The first counterfactual analysis shows that most short-term consumer welfare gains from CGMs go to device users, with the benefits to nonusers being ten times smaller. By revealing previously unobserved attributes and steering demand towards newer drugs, CGMs provide opportunities for physicians to learn 2.7% faster on average about the performance of these drugs in real-life as users share their experiences with the new product with their physician. However, learning about new drugs that are less attractive on the attributes observables thanks to CGMs is slightly slower, -0.7%, driven by a fixed total number of learning opportunities and the limited effect of CGMs on the precision of experience signals that are extrapolated to nonusers. The information generated by CGMs contributed heterogeneously to new drug adoption, ranging from a one percentage point decrease for the entering bioequivalent drug to a four percentage point increase (+16%) in 2021. The drug benefiting most from CGM adoption triggers low overnight glucose levels — a feature more complicated to detect without the device. Turning to negotiated insulin prices, the attributes observed thanks to the technology affect the perceived relative quality of products, which matters for the regulator who maximizes consumer welfare. As a result, the manufacturers of drugs that perform better on these new attributes bargain slightly higher prices, up to +4.4% in 2021. For most products, demand and supply forces reinforce each other, and the overall impact on product-level profits ranges from a 13% decrease to a 23% increase.

The second counterfactual highlights the complementarity between innovations in the medical device and pharmaceutical markets. A comparison of profits from alternative insulin products in environments with and without CGMs indicates that the most profitable pharmaceutical innovation strategy can shift depending on the technological environment in which insulin choices are made. Specifically, when considering the entry of a potential new drug with relatively strong performance on attributes unobservable without CGM technology, its profits would be 10% lower than those of the actual market entrant in an environment without

CGMs. However, in an environment where CGMs are widely used, the profitability of this potential entrant significantly increases, becoming 64% higher than that of the actual market entrant. This increase in profitability comes with a 15% rise in consumer welfare compared to the welfare generated by the current market entrant. This result implies that devices like CGMs, which introduce new observable attributes for evaluating drug performance, may shape future pharmaceutical innovations.

Related literature The primary contribution of this paper is to develop an empirical framework to understand how digital technologies can overcome demand-side information frictions. In this context, it contributes to several strands of the existing literature. Primarily, it builds on research analyzing demand-side learning in pharmaceutical markets. The existing literature emphasizes that information frictions regarding drugs' clinical match values arise upon market entry and initial diagnosis (Coscelli and Shum (2004), Crawford and Shum (2005), Currie and MacLeod (2020)). These frictions decrease as physicians gain direct and indirect experience with the drug — such experience can come from drug manufacturers' detailing, scientific information, clinical trials, etc. (Coscelli and Shum (2004), Zhu (2023), Dickstein (2021), Grennan et al. (2024), Dubois and Tunçel (2021), Alsan et al. (2024)). Building on structural modeling of demand-side learning from experience (Roberts and Urban (1988), Erdem and Keane (1996)), my work is closely related to Crawford and Shum (2005) who focus on treatment experimentation and learning upon initial patient diagnosis. In this paper, I study learning about new products upon market entry, focusing on patients *already* diagnosed with the disease at the time of entry. Physicians learn dynamically about the drugs' clinical match values across their patient population, drawing on experiences of heterogeneous qualities depending on patients' monitoring technology.[4]

When consumers learn dynamically about demand, firms have incentives to adopt forward-looking pricing strategies (Shapiro (1983), Bergemann and Välimäki (2006)). In this context, modeling pricing decisions is inherently complex, leading most empirical literature to focus on demand-side mechanisms while keeping prices fixed in counterfactual scenarios. Exceptions in the pharmaceutical context are Ching (2010a) and Ching (2010b). My paper contributes to this literature by building a tractable framework in which physicians' learning is internalized by the pharmaceutical manufacturers and the regulator when engaging in Nash-bargaining over

[4]Beyond learning, this work adds to the literature on inertia sources in demand (Dubé et al. (2010)). In a chronic condition treatment, I document switching costs contribution separately from learning by leveraging the specificities of my institutional setting.

the price of treatments.

This project also contributes to the literature on the impact of information frictions on market outcomes. On the demand side, previous works document how consumer choice under imperfect information affects product offerings (Brown and Jeon (2024)) and how information provision to consumers can shift the market equilibrium (Jin and Leslie (2003), Handel and Kolstad (2015), Barahona et al. (2023)). On the supply side, incomplete information about the demand curve affects firms' behavior beyond insurance markets (Hitsch (2006), Handel and Misra (2015), Doraszelski et al. (2018)). I focus on the impact of patient-level data to overcome information frictions. In the auto insurance market, consumer-level data generated by monitoring technologies can mitigate information asymmetries (Jin and Vasserman (2021)). Instead, I focus on demand-side information frictions and data-driven insights from a monitoring technology that informs consumers about their effective preferences for alternative products.

I add to the literature on assessing the value of innovative pharmaceutical drugs and medical devices, drawing on the method developed in the empirical industrial organization literature (Trajtenberg (1989), Petrin (2002)). Igami et al. (2024) measure the welfare gains from product and process innovations. In this paper, I link the last two literatures by studying how new information technologies producing insights available for decision-making can enhance the value of pharmaceutical innovation for consumers by guiding physicians' choices.

This work also explores complementarities across related markets and how innovations in one market may shape outcomes in another. While earlier work documents the impact of upstream innovation in vertically related markets (Eizenberg (2014)), Bresnahan and Trajtenberg (1995) highlight that complementary technologies may be necessary to fully exploit the potential of new technologies. My paper emphasizes the complementarity between medical device innovation and the diffusion of new pharmaceuticals in a context where the adoption of the complementary technology is the consumer's choice. I show how innovations in the medical device market alter pharmaceutical product shares and could shape the product offering.[5]

Finally, this paper contributes to the literature on information technology adoption in healthcare. Earlier empirical works focus on the impact of electronic medical records on healthcare costs and hospital productivity (Agha (2014), Dranove et al. (2014), Lee et al. (2013), McCullough et al. (2016)) and, more recently, telemedicine (Zeltzer et al. (2024), Dahlstrand

[5]Previous work by Dranove et al. (2022) highlight that demand shocks can incentivize firms to undertake R&D activities. However, in the context of Medicare Part D, the effect favors follow-up rather than breakthrough innovations.

(2024)) or artificial intelligence (Agarwal et al. (2023)). Handel and Kolstad (2017) highlight the potential for wearable devices to overcome the lack of data on critical outcomes. Yet, to the best of my knowledge, the literature studying how new information technologies impact healthcare markets is scarce. I extend this literature by documenting the channels through which wearable devices generating high-frequency health data affect treatment choices.

The paper proceeds as follows. Section 1.2 provides background information on diabetes treatment in France and describes the data. Section 1.3 documents CGM adoption and its impact on insulin choice. Section 1.4 develops and estimates a demand model for insulin, with some patients using digital devices. Section 1.5 models the pricing decision and estimates the model's primitives. Section 3.6 presents the counterfactual scenarios, and Section 3.7 concludes.

1.2 Context and Data

1.2.1 Diabetes treatment in France

Diabetes is a major chronic condition affecting 1 in 10 adults worldwide.[6] The diseases is characterized by high blood sugar levels which can arise as the pancreas stops producing insulin (Type 1 Diabetes) or as the insulin produced loose its efficacy (Type 2 Diabetes). Both high and low glucose levels can lead to severe complications, including blindness, amputation, stroke, heart attack and kidney failure. The primary goal of diabetes management for the patient (he) and the physician (she) is to stabilize blood glucose within a targeted range.

As of 2021, approximately 22% of diabetic adults in France were dependent on insulin. This project focuses on the choice of long-acting insulin, referred to throughout the paper as *insulin choice*. Long-acting insulin is often combined with short-acting products. While short-acting insulin is injected around meal times to manage food intake, long-acting insulin is designed to stabilize the glucose level over 24 hours. Since these two types of insulin serve different functions, they can be studied independently. Long-acting insulins are primarily differentiated by their theoretical duration of action and must be injected daily.[7] However, the effective duration of each product varies across patients, depending on demographics, time since diagnosis and individual metabolism. Hence, the clinical benefit of a product is heterogeneous across

[6]Retrieved from Diabetes Atlas https://diabetesatlas.org on October 8th, 2024. The prevalence of diabetes among adults is 5.3% in the French population, compared to 10.7% in the U.S.

[7]In May 2024, the European Medicines Agency (EMA) approved the first once-weekly long-acting insulin.

patients.[8] Part of daily diabetes management requires patients to monitor their glucose levels; the objective is to adjust short-acting insulin doses, avoid adverse events, and prevent complications. Before 2017, glucose monitoring relied on disposable strips, representing a significant burden as one must prick his finger for each glucose measure.[9] To evaluate diabetes control, physicians relied on the three-month average glucose level (A1c), glucose measurements from strips and patient-reported adverse events. In particular, the laboratory-measured average was the gold standard for assessing good vs poor diabetes management.

Between 2015 and 2021, significant changes occurred in the insulin product space and the glucose monitoring technologies, impacting the information available to physicians to evaluate treatment efficacy. On the insulin side, before 2016, the set of products available was limited, with Lantus, a 24-hour insulin, accounting for more than 60% of prescriptions. Between 2016 and 2018, four new products entered the market (Figure 1.1). These new insulins include a bioequivalent drug for the 24-hour insulin and three products with an extended duration. These latter target patients for whom the effective duration of the 24-hour product is less than a day and the 42-hour product specifically targets patients with a low glucose level overnight. Prescription drugs are experience goods, meaning that physicians learn about the new insulins' performance outside the controlled environment of clinical trials across a heterogeneous patient population. On the glucose monitoring side, Continuous Glucose Monitors provide glucose readings every 5 to 15 min through a sensor, generating continuous data (Figure 1.21). The technology contrasts with the unique snapshot provided by the disposable strip tests. In France, CGMs became widely used by patients following the Health Technology Agency (HTA) coverage decision in mid-2017. The decrease in daily glucose monitoring burden thanks to CGMs drives a broad and fast device adoption. The coverage decision targeted around 68% of diabetic patients taking long-acting insulin, and the cost of monitoring glucose in France increased by 67% between 2014 and 2023 (Figure 1.22).[10] Indeed, CGMs rely on a disposable sensor that must be replaced every two weeks, making this form of monitoring substantially more expensive than traditional test strips.

CGMs expand the attributes observables for evaluating insulin treatment by providing de-

[8] Adjusting the insulin dosage has limited effects on its duration, as it impacts glucose levels throughout the entire period of action. For instance, increasing the dosage of a drug that lasts for 20 hours will not extend its duration beyond 20 hours but may cause low glucose levels throughout the day. Therefore, choosing the right product for the right patient represents a substantial decision margin.

[9] Regular glucose monitoring involves placing a drop of blood on a disposable test strip, read by a glucose meter.

[10] Guerci et al. (2023) note that despite this restriction, some patients excluded from the coverage were prescribed the technology, leading to coverage extension in June 2023.

tailed measurements of daily glucose profile, including overnight levels. They complement traditional metrics, as the laboratory-measured average glucose does not capture the (i) within-day variation or (ii) day-to-day fluctuations in glucose levels and can obscure important heterogeneity in glucose control (Figure 1.19).[11] By providing insights into the glucose profile, CGMs generate some information that (i) was previously unavailable and (ii) matters when evaluating the performance of insulin therapy. In a market where new products are entering, the impact of digital devices on insulin choice can be twofold: (i) identifying poor patient-product clinical matches and (ii) gathering information about the real-life performance of new drugs. The following patient quote from a medical case study illustrates the first effect.

> *"Last year I found it very helpful to switch to using a [CGM]. [...] I would frequently find that, on the [24-hour] regime, I would experience night-time lows. [...] My consultant and I agreed on a trial of splitting the dose of [24-hour] between morning and evening, but this did not suit me. [...] Following this, my consultant switched me onto [36-hour], which worked much better [...]."* Shields and Sankaranarayanan (2016)

Moreover, Seaquist et al. (2017) highlights the value of CGM information in the context of new insulins' prescriptions.[12] In the remainder of the paper, I will refer to products by their duration of action, '24-hour Biosimilar' for Abasaglar, Type 2 for Xultophy and 'Human' and 'Mix' for 12-hour products.

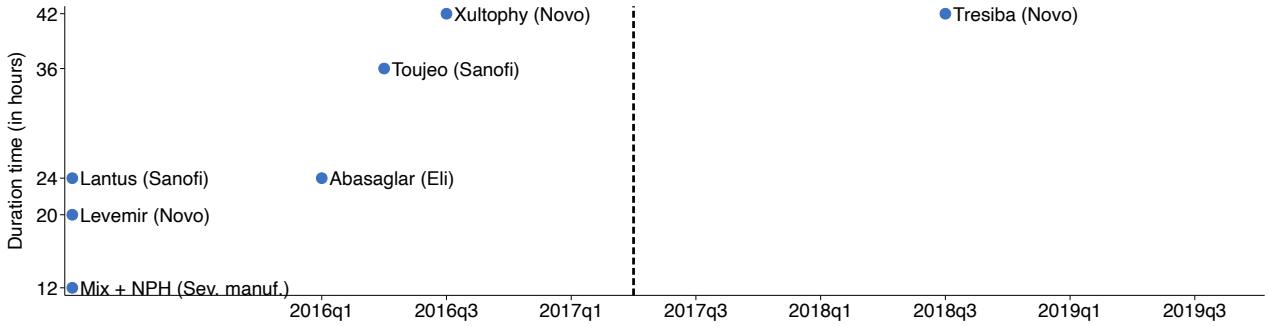
1.2.2 Data

I rely on rich claims data from the French health insurance system. Owing to the centralized universal healthcare system, the data includes all care reimbursements to a given patient and all care prescribed by a physician.

[11] In recent years, concerns about over-treatment have emerged, driven by physicians overestimating the benefit of low average glucose levels at the expense of the risk of low glucose underlying the variance of glucose profile. https://www.reuters.com/investigates/special-report/usa-diabetes-overtreatment/ https://www.medicalnewstoday.com/articles/326063#Millions-of-people-receive-too-much-therapy

[12] "Data from CGM profiling and glycaemic variability studies are providing new and important insights on clinical outcomes with basal insulins in patients with diabetes. These data should enhance confidence in the use of the newer basal insulins in clinical practice by providing physiological context to real-world observations from heterogeneous patient populations." Seaquist et al. (2017)

Figure 1.1: New products entry timeline, 2016-2019



Note: The horizontal axis corresponds to the entry date for new products. The vertical axis corresponds to the theoretical duration of each insulin. Each dot corresponds to one product, except for the 12-hour products. The dashed line corresponds to the reimbursement of the monitoring device.

Insulin prescriptions

The data on insulin prescriptions comes from pharmaceutical claims from 2015 to 2021. For each reimbursement flow, I observe the patient and prescriber IDs, the medical speciality of the prescriber, the date of the prescription, the date of the claim (i.e. the date of the pharmacy visit) and the drug characteristics at the package level. The data is exhaustive of the French population. Hence, it covers all insulin reimbursements to a given patient and all prescriptions written by a physician.[13] The prescription date allows one to account for distinct visits to a particular physician separately from refills.

CGM adoption and attrition

Observing patients adopting and dropping out of continuous glucose monitoring is a crucial analysis component. CGMs are registered medical devices whose coverage was enacted in France in June 2017. CGM use is inferred from pharmaceutical claims data. Like insulin reimbursements, I observe the patient and prescriber IDs, the prescription date, the pharmacy visit date, and the device characteristics. I assume the patient starts wearing the sensor on the day it is claimed at the pharmacy and stops using CGM at the expiration of the last sensor reimbursed to the patient.[14] Mismeasurement in CGM adoption can arise for patients purchasing the technology out-of-pocket (i) before June 2017 or (ii) because they are not eligible

[13]Purchasing insulin requires a medical prescription. I observe all prescriptions filled by the patient. In practice, the prescription is written by the physician and purchased by the patient, leaving scope for nonadherence. I assume that patients always comply with their physician's prescription; hence, I observe all prescriptions written by a physician. Patients under insulin therapy must inject insulin every day, limiting the scope for nonadherence.

[14]The current CGM technologies rely on disposable sensors lasting up to 14 days. I use the duration of each sensor to infer potential attrition from continuous glucose monitoring.

for coverage from the health insurance system.[15] In Appendix 1.9.2, I describe how I recover the effective adoption date for eligible patients who might have adopted before June 2017. The eligibility criteria for CGM coverage include diabetic patients receiving both short- and long-acting insulin daily (basal-bolus therapy). This project focuses on long-acting insulin choice, and patients may rely exclusively on long-acting insulin. Approximately 68% of the population of interest is eligible for CGM coverage. Appendix 1.9.2 provides more details about the risk of mismeasurement for non-eligible patients.

Patient-level demographics and medical conditions

Patient demographics are crucial to my analysis, as diabetes affects a diverse range of individuals, partly due to the two distinct sources of the disease (Section 1.2.1). The data includes the patient's age, gender, and residential area at the municipality level. Low-income individuals are identified as they benefit from free supplement insurance from the government. The medical information includes the type of diabetes and chronic conditions reported in annual patient registries. These files are built from all drugs and care reimbursements to a patient over a year. They report information about anxiety, cancer, cardiovascular disease, dialysis, depression, hypercholesterolemia, hypertension, obesity, etc. Claims data register biological tests' occurrence, but the results are not reported.

Inpatient care/Emergency Room visits

One concern regarding CGM adoption is that patients who adopt the technology have worse diabetes management before the technology is available. To rule out this concern, inpatient care and Emergency Room (ER) visits are used to assess diabetes management before CGM adoption. The data includes all ER visits, regardless of whether they result in an inpatient stay. However, I cannot specifically identify ER visits related to diabetes unless they lead to an inpatient stay, as diagnosis codes are only recorded in those cases.

1.2.3 Sample selection

In this study, I focus on diabetes specialists' prescriptions from 2015 to 2021 for patients who were already taking long-acting insulin before 2016 and were between 18 and 75. This patient-physician sample and time horizon are convenient for several reasons. (i) The period

[15] The first device covered by the health insurance system was approved in the EU in August 2014.

covers the entry of new drugs from 2016 and CGMs coverage from 2017. (ii) Diabetes patients who had used insulin before 2016 were familiar with insulin injection and glucose monitoring before CGMs introduction.[16] (iii) I focus on adults below 75 years old to ensure the patient injects insulin himself. (iv) Diabetes specialists account for only 25% of prescriptions, but General Practitioners (GPs) renew their prescriptions. Hence, they considerably impact insulin prescribing and are likely to make an active choice during a medical appointment.[17] Appendix 1.9.1 provides more details about the data construction and sample selection.

Table 1.11 provides some summary statistics about the incumbent patients followed by diabetes specialists. The final sample focuses on around 330k patients, including 28% of Type I diabetes and 39% of Type II patients using short and long-acting insulin therapy. 68% of the sample was eligible for CGM reimbursement. Patients are, on average, 57, and less than half are women. Around 13% benefited from the low-income complementary insurance during 2015-2022. They live in less favoured areas than the average French individual. Type II patients are more likely to take treatments for other chronic conditions such as hypertension and hypercholesterolemia.

Table 1.12 provides summary statistics for physicians involved in the diabetes treatment for my patient sample. It suggests significant heterogeneity across physicians involved in insulin prescription. Most appointments happen with GPs, who represent 90% of entities and 72% of visits. Yet, the median number of patients per practice is limited. I focus on diabetes specialists and hospital services who see more patients and write more prescriptions at a given practice. While they represent 25% of prescriptions, they account for 74% of treatment switches. Irrespective of their speciality, 58% of physicians faced patients wearing a CGM between 2017 and 2021.

1.3 Descriptive evidence

This section presents motivating evidence that the adoption of CGM by diabetic patients interacts with insulin demand. Patients treated with long-acting insulin widely adopted CGM as a monitoring system following the coverage decision. After adopting the device, patients

[16] The results will not capture CGMs benefits to learn about insulin injection therapy and diagnosis matching upon insulin therapy initiation (Crawford and Shum (2005)). I exclude diabetes patients relying on a pump to inject insulin as they do not rely on long-acting insulin daily.

[17] I further restrict diabetes specialists to the ones prescribing insulin before 2016. By focusing on specialists, I also limit the extent of heterogeneity across physicians, in particular, to control for learning from indirect experience such as detailing by pharmaceutical companies (see Appendix 1.10.2), scientific articles, etc.

are more likely to switch towards new insulin products. The features of the switching behavior are inconsistent with lower switching costs with the digital device. Evidence of physician-level learning about new drugs leaves scope for spillovers. These facts motivate the features of the structural model.

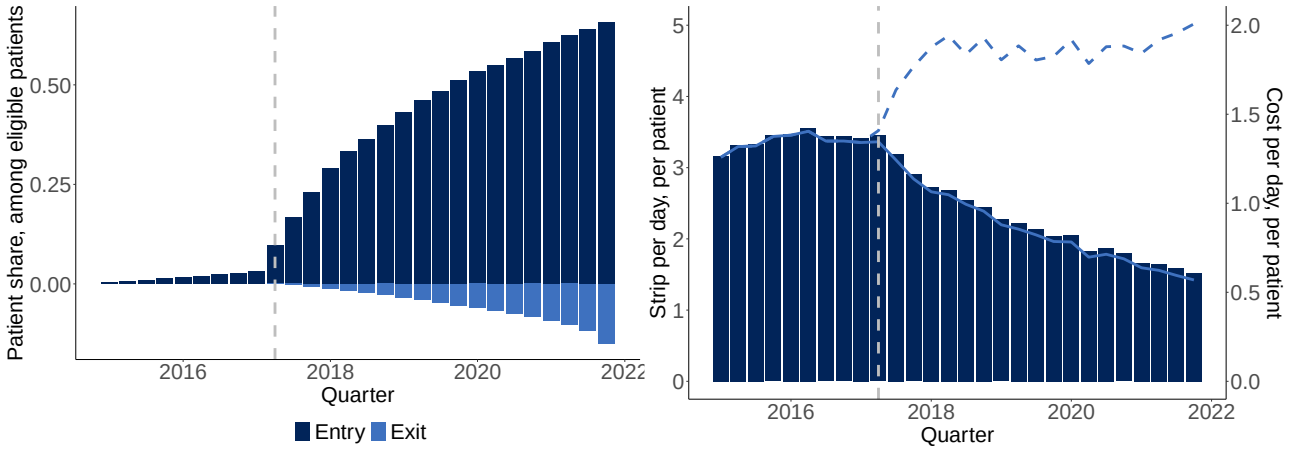
1.3.1 Wide CGM adoption among insulin-treated patients

Following the June 2017 coverage decision, diabetic patients taking insulin widely adopted CGM to replace disposable strips. As shown in Figure 1.2, around 65% of eligible insulin patients ever used the technology, with adoption being staggered. The wide adoption of CGMs decreased the consumption of disposable strips to measure glucose levels while the average cost associated with glucose monitoring increased by 43% (Figure 1.2b). While convenience appears as the primary driver of adoption, I further explain the adoption patterns with respect to patient and physician characteristics. Table 1.1 presents logistic regression results, where the AIC and pseudo-R^2 criteria indicate diabetes type, age and gender are the main patient-level factors influencing CGM adoption. Physician fixed effects account for 31% of the remaining explained variation, whereas environmental factors, other medical conditions and prior diabetes management contribute marginally. In particular, ER visits and diabetes-related hospitalizations in 2015-2016 proxy diabetes control before CGM. A potential concern is that patients opting for continuous glucose monitoring may have been in worse condition before the technology coverage. However, the minimal contribution of hospitalizations pre-2017 in explaining adoption (1.5%) mitigates this concern in the case of CGM uptake.

I also consider whether patients adopted CGM as a result of a worsening of their condition, leading to selection concerns. I study the correlation between CGM adoption and the occurrence of diabetes-related ER visits before the adoption date. For patients adopting CGM, I consider the number of ER visits occurring in a bandwidth of 365 days before the first CGM prescription. The potential adoption date does not exist for patients not adopting the technology. Instead, I consider the maximum number of diabetes-related ER visits the patient faces in 365 days from June 2017 to January 2020.[18] If a worsening of diabetes management drives the technology adoption, ER visits should be positively correlated with CGM adoption. Table 1.2 presents the results, suggesting a negative correlation between the two variables. This evidence indicates that diabetes-related severe adverse events are unlikely to drive technology adoption.

[18]The correlation is estimated for different bandwidths. ER visits after January 2020 are not included due to

Figure 1.2: Adoption among eligible patients



(a) Adoption over time

(b) Strips glucose monitoring, volume and cost

Note: Figure 1.2a presents the stock of patients who ever adopted CGM and dropped out from continuous glucose monitoring over time among eligible patients. Adoption corresponds to the first CGM prescription date. Dropout corresponds to the expiration of the last sensor claimed at the pharmacy. Figure 1.2b represents the average number of strips per patient among eligibles (left axis) and the average value of glucose testing (right axis) over time. The solid line corresponds to the costs of glucose strips; the dashed line includes the cost of strips and CGMs. Patients do not stop consuming glucose strips completetly after adopting a CGM as they may need to confirm symptoms of adverse events.

Table 1.1: Drivers of adoption

Variable	Df	AIC	Pseudo R^2
Model	1898	246,802.1995	0.1346
Removing			
- Demographics	1,887	256,050.62	0.1016
- Specialist FE	35	254,621.24	0.0935
- Chronic conditions	1,886	252,170.08	0.1154
- Glucose strips 2015-2016	1,896	247,784.54	0.1311
- Environment	1,894	247,464.48	0.1322
- ER visits 2015-2016	1,892	247,357.04	0.1326

Note: The model is estimated using a logistic regression on the sample of patients eligible for the technology. Demographics include age, gender and diabetes type. Environmental factors include the deprivation index, low-income complementary insurance and city size. Chronic conditions include 12 diseases (including hypertension, hypercholesterolemia, and obesity; see list in Table 1.11). Glucose strips include the average number of strips in 2015-2016. Hospitalizations account for visits before CGM was available.

the beginning of the Covid-19 crisis.

Table 1.2: CGM adoption and prior ER visits

	(1)		(2)		(3)		(4)	
	coef.	s.e.	coef.	s.e.	coef.	s.e.	coef.	s.e.
ER visits (#)	-0.1241	(0.0035)	-0.1001	(0.0033)	-0.1814	(0.0043)	-0.1166	(0.0021)
Patients characteristics			✓		✓		✓	
$\overline{CGM_i}$	0.6206		0.6206		0.6206		0.6206	
$\overline{ER_i}$	0.1127		0.1127		0.0929		0.4661	
Bandwidth (days)	365		365		180		365	

Note: The model is estimated using a logistic regression on the sample of patients eligible for the technology. Standard errors are clustered at the physician level. Patient characteristics include demographics (age, gender, diabetes type), environmental factors (city size, deprivation index of the living area, low-income individual), chronic conditions (see list in Table 1.11) and glucose strips consumption in 2015-2016. Columns (1) to (3) focus on diabetes-related ER visits leading to inpatient stays. Colum (4) focuses on the number of ER visits leading to an inpatient stay, irrespective of the diagnosis.

1.3.2 CGM adoption and potential patient-insulin mismatch

Given the wide adoption of the device among insulin patients, this section aims to document how the adoption of CGM affects insulin choices. I consider the decision by diabetes specialists to switch insulin treatment, prescribing a product in a given appointment v different from the one used by the patient before the visit. I compare the decision to switch product for patients using CGM to the choice made (i) for similar patients and (ii) for the same patient before he adopts. Considering the decision by the patient i-physician k pair in appointment v,

$$Switch_{ikv} = (\beta_0 + \beta_1 First_{iv})CGM_{iv} + \gamma_1 D_i + \gamma_2 X_{kv} + \lambda_k + \delta_q(v) + \varepsilon_{ikv} \tag{1.1}$$

$Switch_{ikv}$ equals one when the treatment prescribed to patient i by physicians k in appointment v differs from the product patient i was using prior to the appointment, zero otherwise. CGM_{iv} is a dummy variable equal to one when the patient comes to the appointment wearing a CGM, and $First_{iv}$ equals one only for the first appointment by patient i to a specialist after adopting a CGM. D_i includes patient demographics and chronic conditions. X_{kv} proxies the physician's information set during appointment v, counting the number of visits with patients already using new products. λ_k refers to a physician fixed effect and $\delta_{q(v)}$ to a quarter fixed effect. The coefficients of interest are β_0 and β_1. $Switch_{ikv}$, the outcome of interest, may be occasional;[19] hence, I consider heterogeneous effects between the first appointment while wearing the device and subsequent appointments. The model is estimated by OLS. Table 1.3 presents the results.

[19] $Switch_{ikv}$ only equals one on the appointment in which the patient switched insulin. If the change is permanent, the outcome variable only equals one once. Patients can switch back and forth between treatments.

Panel A relies on eligible nonusers as a control group. Panel B is restricted to users before they adopt in which case I include patient fixed effects instead of the physician fixed effects. The parameters suggest a positive correlation between wearing a CGM and changing insulin product in the short run. Considering the coefficients beyond the first appointment, the positive correlation disappears. It suggests that patients adopting the technology are more likely to switch treatment only during the first appointment wearing a CGM. The physician seems to react to the insights generated by CGM by switching their treatment for some patients while deciding to stick to the former treatment for others. This evidence is consistent with a patient-insulin mismatch, which the technology revealed to the physician.[20]

Both switching costs and information frictions can contribute to inertia in insulin choice such that either of these mechanisms could rationalize the behaviour documented in the previous paragraph. Dubé et al. (2010) highlights that the different sources of state dependence in demand lead to differences in pricing incentives. I provide evidence that the switching patterns following the adoption of CGM are inconsistent with a reduction in switching costs in the following paragraph. The structural model will thus focus on representing how CGM can reduce information friction through physicians learning from the insights produced by the device.

First, lower switching costs would increase the probability of switching insulin at any appointment. Table 1.3 provides evidence inconsistent with this feature as the positive correlation vanishes beyond the first appointment. Second, switching costs are particularly salient when considering switches from a branded drug to its biosimilar, as the latter must provide evidence of bio-equivalence before entering the market.[21] Figure 1.3 describes the products involved in treatment switches while wearing the technology, plotting the product used at the time of the appointment (horizontal axis) against the product prescribed (vertical axis) conditional on switching insulin in that appointment. Switches from the 24-hour product towards its biosimilar remain infrequent, making lower switching costs for CGM users unlikely. Patients switch away from the 24-hour branded drug to new products with a longer duration of action. In particular, some patients switch to the 36-hour product in the first appointment following adoption, while the product has been available since 2016, before the CGM coverage decision (Figure 1.3a). In Appendix 1.10.3, I rely on the relative choice between the 24-hour branded drug and its biosimilar to estimate switching costs and rule out lower costs for CGM users.[22] Altogether,

[20] The probability of switching back to a former treatment remains low, even for patients wearing a CGM.

[21] Information frictions can still arise from physicians' uncertainty about the bio-equivalence (Maini et al. (2022)).

[22] To that end, I compare the choice probabilities for existing patients who face both information frictions and

this evidence suggests that CGMs provide insights to physicians about the performance of the current treatment, leading to product switches for some patients.

Table 1.3: Pr(Switching) estimates

	(1)		(2)		(3)	
	coef.	s.e.	coef.	s.e.	coef.	s.e.
A. Control Group: Eligibles						
CGM_{iv}	-0.0446	(0.0008)	-0.0154	(0.0008)	-0.0295	(0.0009)
$CGM_{iv} \times First_{iv}$	0.1217	(0.0012)	0.0659	(0.0012)	0.0667	(0.0012)
$CGM\ User_i$					0.0274	(0.0008)
Physician FE		✓		✓		✓
Quarter FE		✓		✓		✓
Patient demographics				✓		✓
Patient × Physician				✓		✓
Physician information set				✓		✓
B. Control Group: Users						
CGM_{iv}	-0.0651	(0.0014)	-0.0356	(0.0014)	-0.0355	(0.0014)
$CGM_{iv} \times First_{iv}$	0.0751	(0.0014)	0.0478	(0.0014)	0.0478	(0.0014)
Patient FE		✓		✓		✓
Quarter FE		✓		✓		✓
Patient × Physician				✓		✓
Physician information set						✓

Note: The sample is restricted to patients eligible for CGM coverage. Panel A and B include quarter fixed-effect. Panel A contains physician fixed-effects and Panel B patient fixed-effects. 'Patient × Physician' controls for the time since the last interaction between the two.

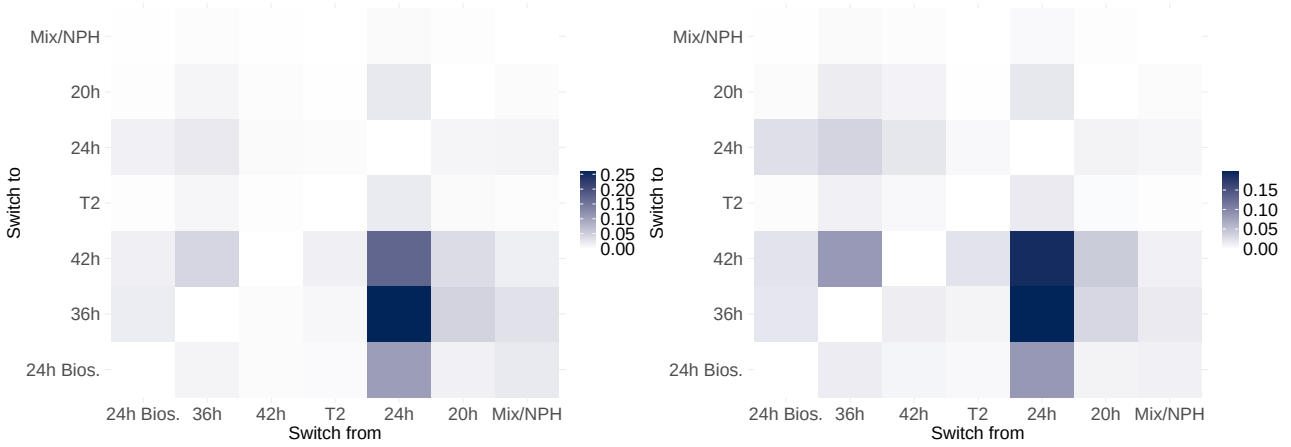
1.3.3 Physician-level learning

The previous section documents treatment switches with the digital device towards newer treatments. Physicians, being less familiar with these products, encounter more learning opportunities thanks to these switches, allowing them to better understand the performance of new drugs. Information externalities arise if physicians apply insights from CGM users to nonusers. I study the correlation between a physician's prescription behavior of new products and her prior product experience to document learning across patients for a given physician.

For each physician k, new product j and quarter q, I compute the product-level prescription share among the patients not already using product j at the time of the appointment. This way, if product j is prescribed to them in quarter q, they were switched to that product by the physician. The number of appointments up to the previous period where the patient already

switching costs, to the first-intention treatment choice made for 'new' insulin patients, who are only subject to information frictions.

Figure 1.3: Switching matrices with the sensor

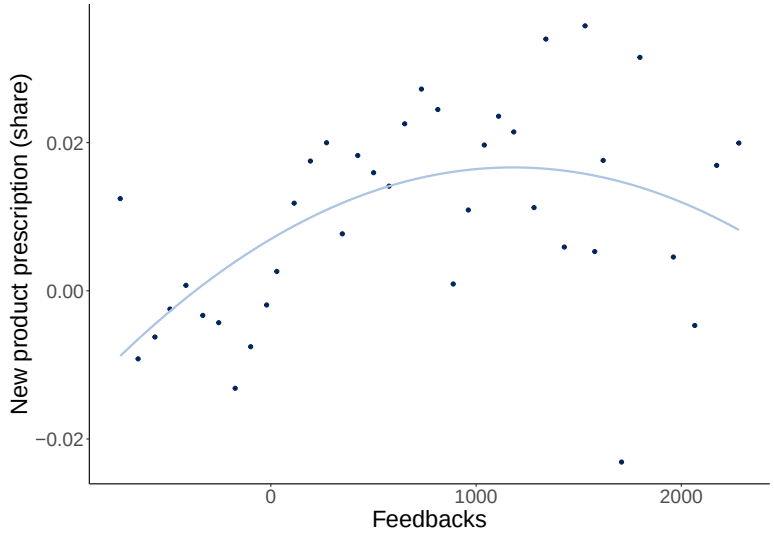


(a) Switches at $CGM_{iv} \times First_{iv}$ (b) Switches when $CGM_{iv} = 1$

Note: Figure 1.3a focuses on the switches occurring at the first appointment after the patient starts to use a CGM, and Figure 1.3b on the switches at any point with a CGM. The horizontal axis corresponds to the product used before the prescription. The vertical axis corresponds to the product prescribed by the physician during that appointment.

used the new product j at the beginning of the visit approximates physicians' information about a given new product j. Indeed, the patient could share information about the drug's performance in his daily life. Figure 1.4 presents a binned scatter plot of a product's information set at the physician level (horizontal axis) and its prescription share to 'new' patients (vertical axis). It suggests a positive correlation between the physician's information set about patients' experience with product j and her propensity to prescribe the product to patients who are not already using the product as they come. This evidence suggests physicians extrapolate information from patients' experiences across individuals, leaving scope for information spillovers between patients adopting a CGM and those not adopting. In Appendix 1.10.4, I document whether the composition of the information set (between patients providing feedback with or without the technology) impacts the strength of the correlation.

Figure 1.4: Physician prescription share and the information set size



Note: The horizontal axis corresponds to the amount of feedback from real-life experience with the new product j received by the physician up to the previous period. The vertical axis corresponds to product j prescription share for patients not using the product before the appointment. For each variable, I consider the residuals from a linear model controlling for physician fixed effects, product-specific quarter fixed effects, and the average demographics of the patient visiting physician k in a given period. The figure focuses on prescriptions by diabetes specialists and excludes the new Type 2 product directed towards a subset of patients. The information set proxy (horizontal axis) is computed at the hospital level for specialists working in the hospital. Figure 1.30 focuses on individual physicians working outside the hospital.

1.4 Insulin demand model

Building on the evidence presented in the previous paragraphs, sections 1.4 and 1.5 develop and estimate a demand and supply model for insulins, considering how digital device information influences prescription drug choices and prices. The model retrieves preferences for insulin products, consumer surplus, and firm profits. Its primitives are estimated using micro-level pharmaceutical claims data, which I use in section 3.6 to quantify the impact of digital device information on the insulin market.

The model explains the behavior of physicians, insulin manufacturers, and the regulator. Insulin manufacturers and the regulator negotiate the price of insulin products. Physicians prescribe an insulin product to their diabetic patients, aggregating up to the demand for each drug. The entry of new drugs and the adoption of CGM are assumed to be exogenous. The timing of the game is as follows. At the beginning of each year, the insulin manufacturers and the regulator agree on the price of each product for the upcoming year. Throughout the year,

each physician (she) faces a sequence of medical appointments with diabetic patients for whom she makes a treatment decision. Physicians face uncertainty about the idiosyncratic match value of a given insulin product for a particular patient, and they learn about clinical match values from patients' experience signals generated with or without a CGM. Insulin manufacturers and the regulator form rational expectations about the upcoming annual demand and learning about the patient-drug match values. When bargaining over drug prices, insulin manufacturers aim to maximize profits from prescription drug sales, while the regulator considers consumer surplus. On the demand side, physicians are assumed to be altruistic and myopic when choosing the insulin product for each patient, leading to a static decision problem at each appointment. Yet, physicians accumulate experience signals over time and can use the information they receive about the real-life performance of products from one patient to inform their decisions for other patients, generating spillovers. This structure enables a two-step approach. In the first step, I model static prescription decisions within each year, aggregating up to the annual demand for each insulin product. In the second step, the aggregate annual demand and consumer surplus for each product are used as inputs to the annual insulin price negotiation between insulin manufacturers and the regulators. This section focuses on the prescription decisions within each year. Section 1.5 presents the price negotiation stage.

1.4.1 Setting

Consider a patient (he), indexed by i, who was diagnosed with diabetes before 2015. He is familiar with insulin injections and glucose monitoring. The patient is followed by a physician k (she), the expert who decides on treatment. At a given physician practice, the flow of medical appointments leading to an insulin prescription is indexed by $v \in \{0, ..., V_k\}$. Physicians differ in the number of patients they see, V_k, and the characteristics of their patient population which are taken as given. The physician prescribes an insulin product $j \in \mathcal{J}_v$ to patient i during a medical appointment. The set of available insulin, $\mathcal{J}_v$, changes across appointments as new products enter the market. The entry of new drugs is assumed to be exogenous. Patients differ in their true clinical match value with product j, denoted $\Theta_{ij} \in \mathbb{R}$, and glucose monitoring technology. The physician faces uncertainty regarding Θ_{ij}. She forms prior beliefs about Θ_{ij} based on clinical trial information and learns about Θ_{ij} through experience signals generated by the direct use of product j. Regarding their glucose monitoring technology, the decision to adopt a CGM and the adoption date are exogenous from the insulin choice perspective. I

denote $a_{iv} \in \{0, 1\}$ the adoption of the digital device, where $a_{iv} = 1$ if the patient uses a CGM, 0 otherwise. The physician observes a_{iv} and the data generated by the glucose sensor when $a_{iv} = 1$. Considering the flow of appointments to physician k, at each visit v,

1. **Patient arrival:** A patient i is coming to the medical appointment for an insulin prescription using a digital or a traditional glucose monitoring device, a_{iv}. The patient's identity, choice of monitoring technology, and set of insulins available are taken as given.

2. **Belief updating:** The physician observes the last product used by patient i until v, $j_{i,v-1}$, and its effect on i's glucose levels. This information generates an experience signal the physician uses to update her beliefs about the clinical benefit of $j_{i,v-1}$ for *any* patient. The physician accumulates experience signals and uses them to learn across patients. Hence, the experience signal enters the information set of physician k at time v, denoted $\mathcal{I}_{kv}$. $\mathcal{I}_{kv}$ represents the stock of information about insulin products' performance received by physician k up to appointment v (included).[23]

3. **Treatment choice:** Given her information set in appointment v, $\mathcal{I}_{kv}$, the physician, k, chooses the treatment $j \in \mathcal{J}_v$ that maximises ones expected payoff.

When selecting the treatment for patient i, the physician faces uncertainty regarding the match value of patient i with product j, Θ_{ij}. The match value is assumed to be independent of a_{iv}, the patient's glucose monitoring technology in a given medical appointment. I assume that the true clinical match value, Θ_{ij}, depends on the preference for the drug's effect on the average glucose level, μ_{ij}, the preference for the drug's effect on the glucose profile, ν_{ij} and that $\Theta_{ij} = \mu_{ij} + \nu_{ij}$.[24] The drug's effect on the glucose profile corresponds to the product's performance at each time of the day. The physician forms beliefs about the true value of each component, μ_{ij} and ν_{ij}, which are updated on the basis of experience signals from her patient population. Upon drugs' entry, the physician's prior beliefs about μ_{ij} may be inaccurate relative to the true value for μ_{ij}. This difference stems from the gap between the average treatment effect from clinical trials and the drug's performance in real-life use among heterogeneous patients.

In the absence of CGMs, the physician receives experience signals and learns about the drug's effect on the average glucose level, μ_{ij}, from direct drug use in their patient population.

[23] The physician receives an experience signal only during a medical appointment linked to an insulin prescription. Given the large number of patients and appointments to a given physician, the empirical specification will impose some restrictions on the extent of learning across patients.

[24] Each argument combines the value of the attribute and the utility weight for that attribute.

Physicians are Bayesians and they learn about μ_{ij} by updating their beliefs using the experience signals received by patients returning to their practice while using product j. I assume that these experience signals are unbiased with respect to the true μ_{ij} but noisy. These signals may be backed up by lab results measuring the three-month average blood glucose level. As a physician accumulates unbiased experience signals, her belief about μ_{ij} converges to its true value. However, the glucose profile of patient i is unavailable without a glucose sensor continuously measuring glucose levels. The physician forms beliefs about the preference for the drug's effect on the glucose profile, ν_{ij}, but she cannot learn about its realization for patient i. I assume $\mathbb{E}(\nu_{ij}|\mathcal{I}_{kv}; a_{iv} = 0) = 0$. As a result, learning about the patient-product clinical match value, Θ_{ij}, cannot be complete without the information produced by CGMs but physicians' belief about Θ_{ij} remains biased.[25]

Glucose sensors measure glucose levels continuously, generating detailed reports about the effectiveness of insulin treatments throughout the day. The analysis of retrospective CGM data by physicians is informative of insulin products' match values. In particular, the benefits from CGM data are two-fold:

1. **Comprehensive measurement of the glucose profile:** The digital device provides detailed information about the glucose profile of patient i. As the drug's effect on the profile is completely observed for the patient wearing the sensor, learning about ν_{ij} is assumed to be complete and immediate once a patient uses the device. The physician observes the realization of ν_{ij} for patient i such that $\mathbb{E}(\nu_{ij}|\mathcal{I}_{kv}; a_{iv} = 1) = \nu_{ij}$. These insights are patient-specific and uninformative for nonusers.

2. **Experience signal about μ_{ij}:** The glucose sensor data generated when the patient was using product j produces an experience signal about μ_{ij}, the preference for the performance of the drug on the average glucose level. This experience signal is also unbiased with respect to μ_{ij} but may be more or less precise than the signal provided if the patient does not use a CGM.

[25] Alternatively, one can assume that physicians learn *partially* about ν_{ij} without a CGM.

1.4.2 Demand

Physicians are incentivized to prescribe the most cost-effective treatment for a given patient. The indirect utility of the physician k-patient i pair when choosing product j is

$$U_{ikjv} = \Theta_{ij} - \alpha p_{jv} + \delta f(age_{jv}) + \varepsilon_{ikjv} \tag{1.2}$$

p_{jv} is the price of product j, age_{jv} the time since drug j is on the market in appointment v and ε_{ikjv} an idiosyncratic preference shock. α measures the price sensitivity of the patient-physician pair, and $f(age_{jv})$ approximates physician's learning about drugs, other than through patients' direct experience. Θ_{ij} is the true clinical match value between patient i and drug j. I assume each physician is myopic when selecting the treatment for patients i in appointment v, maximizing the expected indirect utility from drug j consumption by patient i from appointment v. This assumption limits the extent of experimentation performed by physicians within and across patients. Beyond ethical concerns, this assumption is backed up empirically by the low probability of back and forth between treatments for a given patient. Given the physician information set, $\mathcal{I}_{kv}$, the monitoring device, a_{iv}, the choice of treatment is made according to

$$\max_{j \in \mathcal{J}_v} \mathbb{E}(U_{ikjv}|\mathcal{I}_{kv}; a_{iv}) = \max_{j \in \mathcal{J}_v} \left\{ \mathbb{E}(\Theta_{ij}|\mathcal{I}_{kv}; a_{iv}) - \alpha p_{jv} + \delta f(age_{jv}) + \varepsilon_{ikjv} \right\} \tag{1.3}$$

where $\mathcal{J}_v = \mathcal{J}^{Old} \cup \mathcal{J}_v^{New}$. $\mathcal{J}_v^{New}$ corresponds to the set of insulins entering after 2015 available at the time of appointment v (see Figure 1.1). Given the chronic nature of the disease, there is no outside option, and the physician must prescribe one of the available treatment options. Two disruptions occur simultaneously: new prescription drugs are entering the market, and a new monitoring technology, glucose sensors. Given the information set, $\mathcal{I}_{kv}$, monitoring choice, a_{iv}, and product j, I assume the physician's expectation of the true value for Θ_{ij} is as follows:

$$\mathbb{E}(\Theta_{ij}|\mathcal{I}_{kv}; a_{iv}, j) = \begin{cases} \Theta_{ij} & \text{if } a_{iv} = 1 \text{ and } j \in \mathcal{J}^{Old} \\ \mu_{ij} & \text{if } a_{iv} = 0 \text{ and } j \in \mathcal{J}^{Old} \\ \mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}; a_{iv} = 1) + \nu_{ij} & \text{if } a_{iv} = 1 \text{ and } j \in \mathcal{J}_v^{New} \\ \mathbb{E}(\mu_{ij} + \nu_{ij}|\mathcal{I}_{kv}; a_{iv} = 0) & \text{if } a_{iv} = 0 \text{ and } j \in \mathcal{J}_v^{New} \end{cases} \tag{1.4}$$

In what follows, I provide the intuition for $\mathbb{E}(\Theta_{ij}|\mathcal{I}_{kv}; a_{iv}, j)$ without $(a_{iv} = 0)$ and with

($a_{iv} = 1$) a CGM for old ($j \in \mathcal{J}^{Old}$) and new ($j \in \mathcal{J}_v^{New}$) insulin products.

Absent the digital device, $a_{iv} = 0$

Old treatments ($\mathcal{J}^{Old}$)

The physician has a lot of experience prescribing $j \in \mathcal{J}^{Old}$. Hence, I assume she knows the drug's effect on the average glucose level, μ_{ij}, for each old drug and each existing diabetes patient. This assumption helps address the initial conditions challenge in the physician's information set, $\mathcal{I}_{kv}$, which naturally arises in dynamic settings. However, the physician has no information about the drug's effect on the glucose profile, ν_{ij}, as the exact patient profile i remains unobserved without a glucose sensor. The physician's expectation about ν_{ij} remains equal to her initial belief upon drugs' entry, $\mathbb{E}(\nu_{ij}|\mathcal{I}_{kv}; a_{iv} = 0) = 0$. Without the technology, for product $j \in \mathcal{J}^{Old}$,

$$\mathbb{E}(\Theta_{ij}|\mathcal{I}_{kv}; a_{iv} = 0) = \mu_{ij} + \mathbb{E}(\nu_{ij}|\mathcal{I}_{kv}, a_{iv} = 0) = \mu_{ij} \tag{1.5}$$

μ_{ij} is independent of the physician k and medical appointment v because all physicians have already learned about this value for old drugs. As patients are not using glucose sensors, information about the glucose profile is unavailable to physicians when considering the expected match value between patient i and drug j.

New treatments ($\mathcal{J}_v^{New}$)

For new treatments upon entry, unlike Equation (1.5), the physician has no prior experience, hence imperfect information about the drug's effect on the mean glucose level, μ_{ij}. Upon product entry, at $t = 0$, the physician forms a prior belief about μ_{ij} and ν_{ij}. I follow Erdem and Keane (1996) in assuming beliefs about μ_{ij} are normally distributed such that $\mu_{ij} \sim \mathcal{N}(\mu_{ij}^0, V_{ij}^0)$. As mentioned above, the physician may have nonrational expectations about new drugs' performance on the average glucose level upon entry such that it is possible that $\mu_{ij}^0 \neq \mu_{ij}$. The physician is Bayesian and updates her belief about μ_{ij} on the basis of the patient's experience signals when he returns to her practice at time v using the new product j. She learns *across* patients, using the signal produced by patient i to update the prior belief for any patient i'. These signals, $e_{ii'kj}^v$, are assumed to be unbiased with respect to $\mu_{i'j}$ but noisy. The physician extrapolates the information provided by patient i for patient i'. They are drawn from a normal distribution, $e_{ii'kj}^v \sim \mathcal{N}(\mu_{i'j}, \sigma_{ii'v}^2)$ where $\sigma_{ii'v}^2$ corresponds to the noise of the signal provided by

patient i in appointment v when extrapolated to patient i'. The signal enters the information set of physician k at $v' \geq v$, $\mathcal{I}_{kv'}$. The physician's belief about the preference for the drug's effect on the glucose profile, ν_{ij}, satisfies $\mathbb{E}(\nu_{ij}|\mathcal{I}_{kv}; a_{iv} = 0) = 0$. Without a glucose sensor, she cannot learn about its realization for patient i. Without the digital device, for $j \in \mathcal{J}_v^{New}$,

$$\mathbb{E}(\Theta_{ij}|\mathcal{I}_{kv}; a_{iv} = 0) = \mathbb{E}(\mu_{ij} + \nu_{ij}|\mathcal{I}_{kv}; a_{iv} = 0) = \mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}; a_{iv} = 0) \tag{1.6}$$

where $\mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}; a_{iv} = 0)$ is the mean of physician k's belief about μ_{ij} at time v given that $a_{iv} = 0$. This expectation varies (i) over time as a physician gathers more experience and (ii) across physicians as they see different patient populations. As a physician sees many patients, the size of the information set is large. Section 1.4.3 presents the empirical specification for the learning process and the restrictions keeping the state space tractable.

With the digital device, $a_{iv} = 1$

The digital device generates information about (i) the glucose profile for patient i, leading to immediate learning about ν_{ij} for all drugs for patient i, and (ii) the preference for the drug's performance on the average glucose level, μ_{ij}, for the product used by the patient arriving at the medical appointment v.

Old treatments ($\mathcal{J}^{Old}$)

Thanks to the glucose sensor the realization of ν_{ij} for patient i is observed by the physician such that $\mathbb{E}(\nu_{ij}|\mathcal{I}_{kv}; a_{iv} = 1) = \nu_{ij}$. Given the patient's monitoring device, $a_{iv} = 1$, for drug $j \in \mathcal{J}^{Old}$,

$$\begin{aligned}\mathbb{E}(\Theta_{ij}|\mathcal{I}_{kv}; a_{iv} = 1) &= \mathbb{E}(\mu_{ij} + \nu_{ij}|\mathcal{I}_{kv}, a_{iv} = 1) \\ &= \mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}, a_{iv} = 1) + \mathbb{E}(\nu_{ij}|\mathcal{I}_{kv}; a_{iv} = 1) \\ &= \mu_{ij} + \nu_{ij} = \Theta_{ij}\end{aligned} \tag{1.7}$$

Since there is no uncertainty about μ_{ij} for old drugs, $\mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}, a_{iv} = 1) = \mu_{ij}$. The insights about ν_{ij} from CGMs are patient-specific and uninformative for nonusers who keep facing Equation (1.5).

Owing to the digital device, learning about the performance of old treatments for patient i is complete as the physician now knows Θ_{ij}. Since there is no outside option, pharmaceutical

demand is affected by the monitoring technology only if ν_{ij} differ across alternatives.

Claim 1. *In a mature market ($\mathcal{J}_v = \mathcal{J}^{Old}$), when a digital device generates new insights into the patient-product clinical match value, if there exist at least two $j, j' \in \mathcal{J}$, such that $\nu_{ij} \neq \nu_{ij'}$ then* $\mathbb{E}(\Theta_{ij}|\mathcal{I}_{kv}, a_{iv} = 1) - \mathbb{E}(\Theta_{ij'}|\mathcal{I}_{kv}, a_{iv} = 1) \neq \mathbb{E}(\Theta_{ij}|\mathcal{I}_{kv}, a_{iv} = 0) - \mathbb{E}(\Theta_{ij'}|\mathcal{I}_{kv}, a_{iv} = 0)$.

Despite the lack of product entry, the digital device can affect the insulin demand of device users. This impact is driven by the availability of the patient's glucose profile, a newly observable attribute that makes certain treatments more appropriate. These insights are assumed to be patient-specific. Hence, the magnitude of this effect at the market level strongly relies on device adoption. When insights from glucose sensor data do not emphasize differences across products, for example, when $\nu_{ij} = \nu_i \ \forall j \in \mathcal{J}^{Old}$, demand for existing pharmaceutical alternatives remains unaffected.

New treatments ($\mathcal{J}_v^{New}$)

For new treatments, the impact of CGM data is twofold. First, it allows the effect of treatments on the glucose profile, ν_{ij}, to be inferred. Second, it can affect the quality of the experience signal about μ_{ij} when the patient returns to his physician while using product j, compared to the signal provided without a glucose sensor. For new treatments $j \in \mathcal{J}_v^{New}$,

$$\mathbb{E}(\Theta_{ij}|\mathcal{I}_{kv}, a_{iv} = 1) = \mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}, a_{iv} = 1) + \nu_{ij}$$

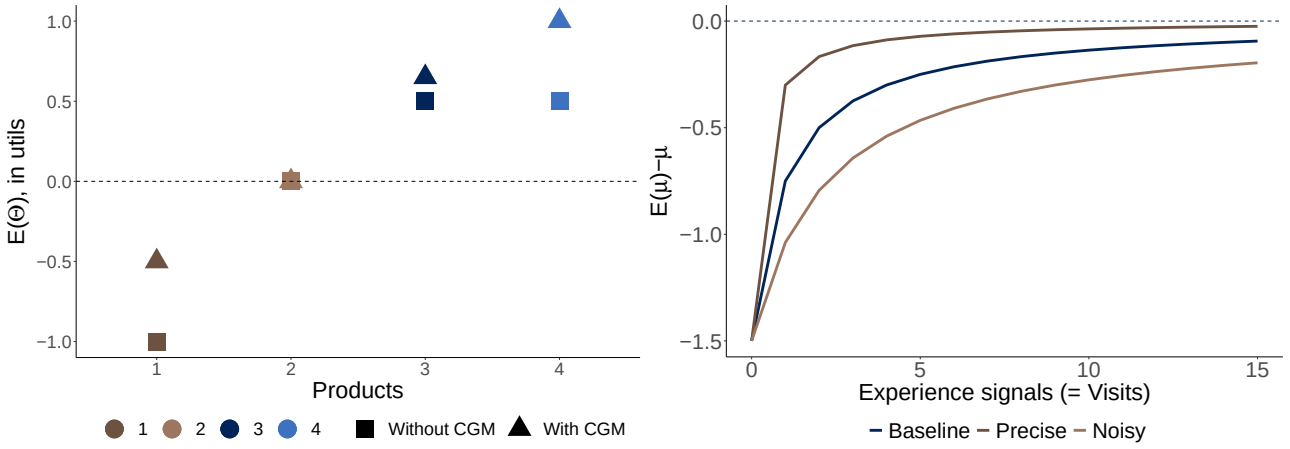
where $\mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}, a_{iv} = 1)$ can differ from $\mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}, a_{iv} = 0)$ in Equation (1.6) for the product the patient was using when coming to the physician visit at time v. The experience signal from a patient returning to his physician using the new product j while wearing the sensor might differ from the experience signal provided without glucose sensor data. I assume signals with and without the device are also normally distributed and unbiased with respect to the true value for μ_{ij} but allow the precision of the experience signal to differ depending on the glucose monitoring technology used by the patient at the time of the appointment. Indeed, observing the performance of a new drug j for a patient i using the digital device, $a_{iv} = 1$, might provide information about μ_{ij} in a different fashion than had the patient come without the technology $a_{iv} = 0$. Whether this experience signal is more, less, or equally informative about μ_{ij} than signals received when the patient is not wearing a CGM remains an empirical question. I do not assume the glucose sensor generates complete learning about μ_{ij}. However, if the insights

from CGM provide precise information about the performance of drug j on the average glucose level for patient i, μ_{ij}, the noise of the experience signal received from glucose sensor will be very small and the belief of physician k after receiving CGM information, $\mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}, a_{iv} = 1)$, will be close to its true value and precise.

Summary Figure 1.5 summarizes the two effects of CGMs on physicians' expectations about the true clinical match value, Θ_{ij}. First, Figure 1.5a presents the impact of comprehensive glucose profile measurement when considering alternative insulin products for a patient, setting aside the dynamic learning about new drugs. The vertical axis corresponds to the physician's expected clinical match values, $\mathbb{E}(\Theta_{ij})$, both with (triangle) and without (square) the device, across various insulin products represented on the horizontal axis. This figure emphasizes how the insights generated by the CGM about the glucose profile can influence the physician's expected match value between patient i and product j by correcting the initial bias in physicians' beliefs about Θ_{ij}. As presented in Equation (1.4), $\mathbb{E}(\Theta_{ij})$, the physician's expectation about the match value between patient i and product j, equals μ_{ij} without the glucose sensor, as the glucose profile for patient i with drug j remains unobserved, and $\mu_{ij}+\nu_{ij}$ with the sensor. The expected match values without a CGM are represented by squares in Figure 1.5a. The physician is indifferent between prescribing products 3 and 4 for patient i, *ceteris paribus*, as their perceived match values are identical. With a glucose sensor, the physician observes the performance of each insulin product j on patient i's glucose profile, ν_{ij}, in addition to μ_{ij}, which makes the expected match value, represented by triangles, now equal to the true value, Θ_{ij}. With information about the glucose profile, the expected match value from product 4 dominates the one from product 3, and the gap between products 1 and 2 narrows. In this example, the technology increases the *perceived* differentiation between products 2, 3, and 4, whereas it decreases between products 1 and 2. If patient i was prescribed product 3 before using a CGM, the insights generated by the device may induce a treatment switch from product 3 to product 4. Given that the true match value between patient i and product j, Θ_{ij}, is independent of the glucose monitoring technology, the glucose sensor only affects the physicians' *perception* of the match values, not the *true* underlying differentiation. Second, Figure 1.5b focuses on the learning dynamics of μ_{ij} for new drugs. It illustrates the impact of experience signals generated by patients who return to their physicians with new drug j while wearing a CGM during their appointment v. CGMs affect the noise of the experience signal. In particular,

the figure represents the evolution of the physician's expectation about the preference for the drug's effect on the average glucose level, μ_{ij}, as she accumulates information. The horizontal axis represents the number of experience signals received, while the vertical axis represents the difference between the physician's expectation of μ_{ij} and its true value. Assuming the data generated by glucose sensors affects the precision of each experience signal, the figure represents the evolution of the physician's belief with (brown lines) or without CGM (navy line). If experience signals generated by CGMs are more precise (dark brown) than traditional signals (navy line), the gap between the physician's expectation, $\mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}, a_{iv})$, and the true value of μ_{ij} is smaller, for a given number of experience signals received (horizontal axis). Hence, the physician can learn faster about μ_{ij} thanks to CGMs. The figure also illustrates that the belief is immediately near the true match value after receiving a precise signal.

Figure 1.5: Impact of CGMs on the perceived match value



(a) Comprehensive glucose profile

(b) Experience signal about μ_{ij}

Note: Figure 1.5a plots the physician's expectation about Θ_{ij} (vertical axis) for different products (horizontal axis), with (▲) and without (■) insights from CGMs. This example abstracts away from the dynamic learning about μ_{ij} for new drugs. Figure 1.5b shows the evolution of a physician's belief about μ_{ij} (vertical axis) as she receives more experience signals (horizontal axis) from patients who return to her practice while using new product j. One feedback corresponds to one visit where the patient shares his experience with product j. I assume that physicians form normally distributed priors about μ_{ij} at $v = 0$ and update their beliefs from normally distributed signals received from patients using product j. The signals are unbiased with respect to the true value, μ_{ij} but noisy, denoting the variance of the signal from patient i' for patient i, $\sigma^2_{i'iv'}$. Here, a physician receives either precise ($\sigma_{i'iv'} = 1$) or noisy ($\sigma_{i'iv'} = 3$) signals compared with the baseline ($\sigma_{i'iv'} = 2$). For exhibition purposes, in this simple example, the noise is the same for signals received from $i' \neq i$.

Consumer welfare under imperfect information In this setting, there exists a discrepancy between the decision utility and that experienced by the patient. Similar concerns arise when the expected utility anticipated at the time of the decision-making and the utility under rational expectations differ (Brown and Jeon (2024)). Uncertainty and incomplete learning distort the decision utility, hence the demand for insulin, whereas the underlying experience

utility relies on Θ_{ij}, the true clinical match value. The experience utility from patient i using product j is

$$U_{ikjv} = \Theta_{ij} - \alpha p_{jv} + \delta f(age_{jv}) + \varepsilon_{ikjv} = \mathbb{E}(U_{ikjv}|\mathcal{I}_{kv}; a_{iv}) - \Big[\mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}, a_{iv}) - \mu_{ij} + \nu_{ij}(a_{iv} - 1)\Big]$$

For old products, $\mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}; a_{iv}) = \mu_{ij}$. Following Dubois et al. (2018) and denoting $j^* = \arg\max_j \mathbb{E}(U_{ikjv}|\mathcal{I}_{kv}; a_{iv})$, the expected experienced utility for patient i from the choice made on his behalf by physician k at v is as follows:

$$\begin{aligned} W_{ikv} &= \mathbb{E}_\varepsilon[U_{ikj^*v}] \\ &= \mathbb{E}_\varepsilon\Big[\max_j\{\mathbb{E}(U_{ikjv}|\mathcal{I}_{kv}; a_{iv})\}\Big] - \mathbb{E}_\varepsilon\Big[\mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}; a_{iv}) - \mu_{ij} + \nu_{ij}(a_{iv} - 1)\Big] \end{aligned} \tag{1.8}$$

where the second term accounts for the difference between the realized and expected clinical match value.

Discussion

This framework relies on several assumptions regarding physician and patient behavior. First, the patient-product clinical match value is not affected by the digital device. It assumes that the information generated by CGMs does not affect the effect of insulin j on the average or profile of glucose levels for patient i.[26]

Second, focusing on patients diagnosed before new product entry, it is assumed that physicians know the preference for the drug's effect on the average glucose level, μ_{ij}, for these patients among old alternatives. This feature circumvents the initial conditions problem that arises in a dynamic setting.

Third, physicians are altruistic and myopic. They are altruistic as they choose the treatment to maximize a weighted sum of the expected patient-level clinical benefit and treatment prices. As physicians' altruism and price sensitivity cannot be separately identified, a low α can be driven by a low sensitivity to price and a high level of altruism. Physicians are also assumed to be myopic, which excludes exploration by physicians who could prescribe a less valuable treatment to learn about the clinical match value of new drugs. This assumption is motivated in the data by the empirical probability of returning to a former treatment which remains low,

[26] Under this assumption, the average and profile of the glucose levels for patient i can still be affected by the technology owing to the better fine-tuning of short-acting insulin around mealtime. Yet, the effect must be independent of the long-acting insulin product choice, j.

even for patients using the digital device. It suggests physicians do not exploit patients with digital devices to collect valuable information about treatments at the expense of a dominated choice for the patient himself. Physicians may be reluctant to prescribe treatments for which the outcome is more uncertain. I approximate this feature in the econometric model by allowing for pessimistic initial priors, capturing the reluctance to switch patients to new treatments.

1.4.3 Empirical model and identification

This section presents the empirical details of the theoretical framework developed in the previous section. Following the last section, the expected indirect utility driving the choice of insulin j by physician k for patient i coming at v is as follows.

$$\mathbb{E}(U_{ikjv}|\mathcal{I}_{kv}; a_{iv}) = \mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}; a_{iv}) + \nu_{ij}a_{iv} - \alpha p_{jv} + \delta f(age_{jv}) + \varepsilon_{ikjv} \tag{1.9}$$

where p_{jv} is the daily price of insulin j at time v, age_{jv}, is the time since product j became available, enters quadratically, $\mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}; a_{iv})+\nu_{ij}a_{iv}$ is the expected patient-drug clinical match value, following Equation 1.4, and ε_{ikjv} is an idiosyncratic shock, unobserved to the econometrician, iid and following a Type I EV distribution. age_{jv} approximates physician learning from indirect experiences such as word of mouth, scientific conferences/articles or pharmaceutical detailing. Physician learning from these alternative sources is assumed to be constant across physicians and product at a given age. The next paragraphs discuss the restrictions imposed on the clinical match value heterogeneity and the parametric restrictions on the learning process about μ_{ij}. Next, I discuss the parameters identification and the potential threats to identification arising from endogenous digital device adoption. Finally, I conclude this section with an overview of the estimation procedure.

Heterogeneity in patient-drug clinical match value

The patient-insulin clinical match value, Θ_{ij}, varies from one patient to another on the basis of physiologic and metabolic factors. Physicians learn dynamically about the preference for the drug's effect on the average glucose level, μ_{ij}, from patient's direct experiences, summarized in the physician's information set in a given appointment, $\mathcal{I}_{kv}$. As physicians see many patients and extrapolate experience signals across patients, the dimension of $\mathcal{I}_{kv}$ is large. To accommodate the heterogeneity in μ_{ij} while maintaining the model tractability, I assume that

patients can be classified, *ex-ante*, into N distinct groups, $n \in \{1, ..., N\}$. Within a cluster, n, the preference for the drug's effect on the average glucose level, μ_{ij}, is constant and denoted μ_{nj}. Physicians know the group to which each existing patient belongs and form beliefs about μ_{nj} upon drug j's entry.[27] Physicians' initial belief about the performance of drug j can be inaccurate such that $\mathbb{E}(\mu_{nj}|\mathcal{I}_{k0}; a_{iv}) = \mu^0_{ikj} \neq \mu_{nj}$. The belief about μ_{nj} evolves as the physician gathers more direct patient experience such that $\mathbb{E}(\mu_{nj}|\mathcal{I}_{kv}; a_{iv})$ varies over time. This value varies across physicians as they encounter patients with different experiences, as summarized in $\mathcal{I}_{kv}$. For bioequivalent products, μ_{nj} is assumed to be equal as the drug's regulatory approval requires evidence of equivalence. Nevertheless, physicians might hold inaccurate beliefs regarding bioequivalence, indicating that the learning process also pertains to the entering biosimilar. As highlighted by Maini et al. (2022), physicians' uncertainty about equivalence may drive slow uptake in prescription shares for biosimilars. Paragraph 1.4.3 below details the specifications for beliefs and signals.

The classification is performed in two steps. First, I categorize patients into three groups based on their diabetes type and insulin therapy. Specifically, I distinguish between 'Type I diabetic patients', 'Type II diabetic patients using both long and short-acting insulins', and 'Type II diabetic patients using only long-acting insulin'. Second, I apply a k-means algorithm within each group to classify patients into two to three subgroups, resulting in a total of 7 clusters. The variables used to create the clusters include patients' demographics, environment, chronic conditions, diabetes management, and health status prior to the availability of continuous glucose monitors. Within a cluster, n, the preference for the drug's effect on the average glucose level, μ_{nj}, is represented by a group-specific product fixed effect. This flexible parameterization allows preferences for specific insulin products to vary across groups on the basis of factors observed by the physician during treatment decisions but not captured in the data.

The preference for drug j's effect on the glucose profile for patient i, ν_{ij}, is proxied by the patient's sensitivity to the insulin duration of action, d_j. In my empirical specification, the digital device reveals the patient's true sensitivity to the duration of action of the insulin products. To accommodate for patient-level heterogeneity and nonlinearities, $\nu_{ij} = \beta_1(x_i)d_j + \beta_2(x_i)d_j^2$ where x_i includes observable demographics and chronic conditions.

[27] For newly diagnosed patients, the physician may not be able to identify i's type directly, leading to diagnosis matching, experimentation, and exploration (Crawford and Shum (2005)). The structural estimation of the model avoids these considerations by focusing on existing diabetes patients.

New products' learning dynamics

Physicians are learning dynamically about the performance of insulin products in real-life patients. Upon market entry, the physician forms beliefs about the preference for drug j's performance on the average and profile of glucose levels for patient i, μ_{ij}, and ν_{ij}. The physician learns about each component with or without a glucose sensor, as presented in sections 1.4.1 and 1.4.2. The μ_{ij} parameter is constant across patients within a cluster n. Hence, a physician learns about μ_{nj} if patient i belongs to group n. I assume that patients in the same cluster n, followed by the same physician k, are subject to the same prior about μ_{nj}. Learning about μ_{nj} occurs at the cluster level. Each physician's prior belief about μ_{nj} and ν_{ij} is summarized by the following distributions:

$$\begin{aligned} &\mu_{nj} \sim \mathcal{N}(\mu_j^0, V_j^0) \\ &\nu_{ij} \sim F(\nu_{ij}) \text{ where } \mathbb{E}(\nu_{ij}|\mathcal{I}_{kv}, a_{iv} = 0) = 0 \end{aligned} \tag{1.10}$$

where $F(.)$ is the cumulative distribution function of the beliefs about ν_{ij}. Each physician holds a normally distributed prior belief about μ_{nj}, characterized by an initial mean μ_j^0 and a precision V_j^0 upon market entry. The initial prior is assumed to be constant across physicians.[28] Physicians are Bayesian and learn about μ_{nj} from the experience signals received from patients who return to their physicians using product j, regardless of whether they are wearing a glucose sensor. I assume that physicians do not learn across clusters, which limits the scope of information spillovers. Consider a patient i in group n, previously prescribed insulin product $j \in \mathcal{J}_v^{New}$, who visits physician k during medical appointment v, with or without a CGM, a_{iv}. A medical appointment generates an experience signal, which is assumed to be distributed according to the following normal distribution.

$$e_{ikj}^v \sim \mathcal{N}(\mu_{nj}, (\sigma_0 + \sigma_1 a_{iv})^2) \tag{1.11}$$

In particular, the experience signals, e_{ikj}^v, are normally distributed and unbiased but noisy. The signal's precision depends on the technology used. Given the normally distributed beliefs and signals, physician k's posterior belief about product j, $\mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}; a_{iv}) \equiv \mu_{nkj}^v$, after receiving i's

[28] Initial prior beliefs could be influenced by the sources of indirect learning, such as clinical trials, medical conferences, and detailing. The model is estimated using a subset of diabetes specialists working outside the hospital. Hence, they are assumed to receive similar information about new products upon entry. In Figure 1.29, I provide evidence that most diabetes specialists interact with the three pharmaceutical companies. The interactions occur on average once per year.

experience, e^v_{ikj}, and given the prior belief, μ^{v-1}_{nkj}, is given by:

$$\mu^v_{nkj} = \begin{cases} \mu^{v-1}_{nkj} \frac{\sigma_0^2}{\sigma_0^2+V^{v-1}_{nkj}} + e^v_{ikj} \frac{V^{v-1}_{nkj}}{\sigma_0^2+V^{v-1}_{nkj}} & \text{if } a_{iv}=0 \\ \mu^{v-1}_{nkj} \frac{(\sigma_0+\sigma_1)^2}{(\sigma_0+\sigma_1)^2+V^{v-1}_{nkj}} + e^v_{ikj} \frac{V^{v-1}_{nkj}}{(\sigma_0+\sigma_1)^2+V^{v-1}_{nkj}} & \text{if } a_{iv}=1 \end{cases} \tag{1.12}$$

The signal's noise when the patient is using a CGM ($a_{iv}=1$) influences how quickly physicians learn from CGMs. If σ_1 is high, the insights from glucose sensors provide an uninformative signal about μ_{nj}, but physicians cannot be misled about μ_{nj} by the CGM readings. Whether signals from CGMs are more, less, or equally informative about μ_{nj} compared to signals from traditional tools depends on the sign of σ_1. Figure 1.5b illustrates the effect of signal precision on the belief over the number of signals.

Discussion The learning framework relies on several assumptions and presents several restrictions. First, learning from experience across physicians is restricted, approximated by the product's time on the market. Beliefs are specific to each physician. The only channel for information sharing between two physicians arises from patients who may consult multiple physicians.[29] Second, there is no learning across different patient types or time discounting for old signals. I assume that physicians retain past information received from patients. Finally, direct experience serves as the only source of physician-specific learning. Physicians are considered homogeneous regarding indirect sources of information, such as detailing, and they are equally likely to prescribe new drugs after conditioning on experience. By focusing on diabetes specialists practicing outside the hospital, there is minimal variation in the physicians' likelihood of being detailed by insulin manufacturers. In Appendix 1.29, I provide suggestive evidence that each insulin manufacturer has detailed the majority of diabetes specialists since at least 2014.

Identification

The set of parameters to estimate includes the price sensitivity parameter and the drug's age coefficient, α and δ, the match value components, μ_{nj}, $\beta_1(x_i)$ and $\beta_2(x_i)$, and the dynamic learning parameters, μ^0_j, V^0_j, σ_0 and σ_1. I assume that the price of each drug and the timing of its entry are exogenous. The endogeneity of insulin prices is unlikely in this context since drug prices are set at the national level, and the empirical specification includes product fixed effects.

[29] In Figure 1.28, I provide descriptive evidence about the limited practice size for diabetes specialists working outside the hospital context. This limits concerns about spillovers across physicians working within the same practice.

The patient and visit-specific taste shock, ε_{ikjv}, is assumed to be uncorrelated with the price of each insulin p_{jv}. The timing of the new drug's entry is also considered to be independent of individual-level taste shocks.

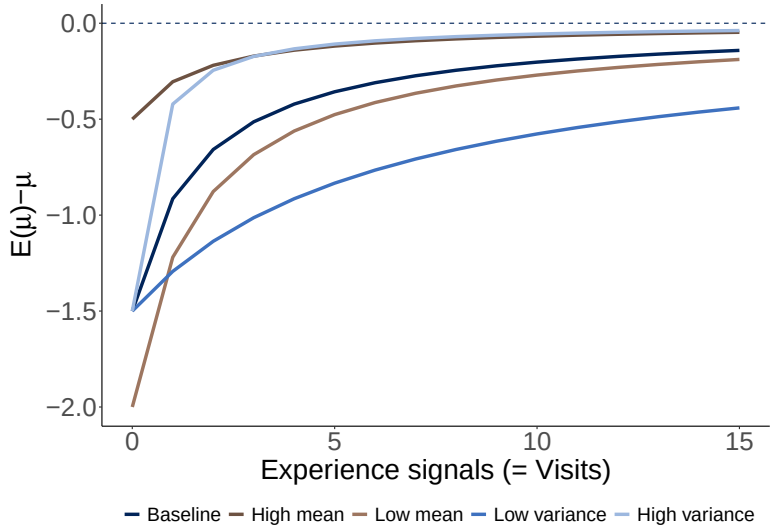
The match value parameters, μ_{nj}, $\beta_1(x_i)$ and $\beta_2(x_i)$, are identified from the choice probabilities of each insulin product j for different patient groups. In particular, the preference for the drug's effect on the average glucose level, μ_{nj}, is identified from within-group choice probabilities for patients without the technology. There is no outside option. Therefore, I normalize μ_{nj} for the 24-hour product to zero for each patient group. The sensitivity to insulin duration parameters, $\beta_1(x_i)$ and $\beta_2(x_i)$, are identified from the choice probabilities for patients using the technology and their deviation compared with nonusers within a group. The causal impact of the digital device on insulin choice can be estimated if ε_{ikjv} is uncorrelated with CGM adoption, a_{iv}. Paragraph 1.4.3 discusses this assumption.

The remaining parameters drive the evolution of physicians' beliefs as they gain experience and encompass the initial prior mean and variance, μ_j^0 and V_j^0, along with the experience signals noise parameters, σ_0 and σ_1. These parameters are identified by leveraging the sequence of prescriptions written by a physician for patients in the same cluster as she gathers more experience signals from any patient in that group. Figure 1.6 presents the intuition for identifying the remaining parameters. The extent of prescriptions made without experience identifies the prior mean. The product-specific variance parameters are identified by the increase in the propensity to prescribe a specific product to any patient in the cluster as the physician acquires experience signals.

Endogenous CGM adoption

I assume that the adoption of the digital device is exogenous to the choice of long-acting insulin, and identifying the causal effect of the digital device on insulin demand relies on the taste shock, ε_{ikjv}, being uncorrelated with CGM adoption, a_{iv}. One potential threat to identification arises from unobserved patient-level characteristics that may vary over time. First, as noted in Section 1.4.2, only unobserved components that lead to differences in product-level clinical match value are relevant. Any individual and time-specific component that does not vary across drug choice alternatives cancels out in the logit specification. Second, the cluster-level product fixed effect accounts for product-level unobserved heterogeneity across patient groups. As a result, the remaining threats lie in within-group product-specific unobservables that correlate

Figure 1.6: Identification of beliefs



Note: This figure illustrates the evolution of the physician's beliefs about μ_{nj} (vertical axis) as she accumulates experience signals (horizontal axis) from patients who return to her practice while using the new product j. Each piece of feedback corresponds to one visit where the patient shares their experience with product j. Physicians are assumed to form normally distributed priors regarding the drug's effect on the average glucose level, μ_{nj} at $t = 0$, and update their beliefs from the Bayes rule based on normally distributed signals. Here, $\mu_{nj} = 0$, and the figure depicts alternative updating of beliefs over time for varying prior means (in blue) and varying prior variances (in brown).

with adoption, which I cannot address.

Another threat to the validity of the estimates arises from physicians strategically offering CGMs to diabetic patients whom they suspect have poor glucose control with their current long-acting insulin. Physicians who are more likely to encounter patients using CGMs would also be more inclined to modify their patients' insulin treatment after adoption. Section 1.3.1 suggests that physicians differ in their likelihood of facing CGM patients (Table 1.1). Concerns about reverse causality are alleviated by three facts. First, the data indicates that most diabetic patients who adopt continuous glucose monitoring do so during their initial visits to a diabetes specialist following the insurance coverage decision. Suppose physicians were gradually becoming acquainted with the benefits of CGM insights for long-acting insulin choices. In that case, one might expect CGM adoption to occur more slowly and in a more staggered manner after the insurance coverage decision. Furthermore, dropout rates from continuous glucose monitoring remain low according to the data, suggesting that patients do not spontaneously use a glucose sensor solely to adjust their insulin prescriptions. Second, as highlighted in Section 1.3.1, hospitalizations prior to CGM adoption are not positively correlated with the adoption

decision. Third, in Figure 1.31, I display the physician-level propensity to treat CGM patients alongside their propensity to switch the insulin treatment of CGM patients during the first appointment after patient adoption. The figure indicates that there is no correlation between the two dimensions.

Estimation

The likelihood of observing choice j by physician k for patient i, $\mathcal{L}_{ikjv}$, using a_{iv}, in appointment v is given by

$$\begin{aligned}\mathcal{L}_{ikjv} &= \Pr(\mathbb{E}(U_{ikjv}|\mathcal{I}_{kv}; a_{iv}) > \mathbb{E}(U_{ikj'v}|\mathcal{I}_{kv}; a_{iv}) \ \forall j' \neq j) \\ &= \frac{\exp(u_{ikjv}(p_{jv}, d_j, a_{iv}|\mathcal{I}_{kv}))}{\sum_{\forall j'} \exp(u_{ikj'v}(p_{j'v}, d_{j'}, a_{iv}|\mathcal{I}_{kv}))}\end{aligned} \tag{1.13}$$

where

$$\begin{aligned}\mathbb{E}(U_{ikjv}|\mathcal{I}_{kv}; a_{iv}) &= \mathbb{E}(\mu_{ij}|\mathcal{I}_{kv}; a_{iv}) + \Big(\beta_1(x_i)d_j + \beta_2(x_i)d_j^2\Big)a_{iv} - \alpha p_{jv} + \delta f(age_{jv}) + \varepsilon_{ikjv} \\ &= u_{ikjv}(p_{jv}, d_j, a_{iv}|\mathcal{I}_{kv}) + \varepsilon_{ikjv}\end{aligned} \tag{1.14}$$

While physicians observe the realization of the signal, e^v_{ikj}, the econometrician does not. Therefore, the likelihood for a given sequence of choices made by physician k must integrate over the distribution of unobserved signals. Given that $y^v_{ikj} = 1$ for the chosen alternative and 0 otherwise, the individual likelihood for physician k is expressed as

$$\mathcal{L}_k = \int_{-\infty}^{\infty} \left[\prod_{v=0}^{V^k} \prod_{\forall j} \left(\frac{\exp(u_{ikjv}(p_{jv}, d_j, a_{iv}|\mathcal{I}_{kv}))}{\sum_{\forall j'} \exp(u_{ikj'v}(p_{j'v}, d_{j'}, a_{iv}|\mathcal{I}_{kv}))} \right)^{y^v_{ikj}} |\vec{e}_k \right] dF(\vec{e}_k) \tag{1.15}$$

where $\vec{e}_k = \{e^1_{ikj}, ..., e^{V_k}_{ikj}\}$, is the vector of signals observed by physician k. Without spillovers within a physician, across patient types, n, the likelihood function can be written at the cluster-physician level. I simulate $M = 200$ Halton draws from a normal distribution to approximate the integral.[30] Given the number of physicians, $\mathcal{K}$, the demand model parameters are estimated via simulated maximum likelihood by taking the simulated log-likelihood of the sample.

$$\log L = \frac{1}{\mathcal{K}} \sum_{\forall k} \log \frac{1}{M} \sum_{\forall m} \left[\prod_{v=0}^{V^k} \prod_{\forall j} \left(\frac{\exp(u_{ikjv}(p_{jv}, d_j, a_{iv}|\mathcal{I}_{kv}))}{\sum_{\forall j'} \exp(u_{ikj'v}(p_{j'v}, d_{j'}, a_{iv}|\mathcal{I}_{kv}))} \right)^{y_{ikjv}} |\vec{e}^m_k \right] \tag{1.16}$$

[30] The draws are generated once and used across iterations.

1.4.4 Results

Estimates

Table 1.4 shows the estimated parameters, excluding the match value components, μ_{nj}, $\beta_1(x_i)$, and $\beta_2(x_i)$. The price sensitivity of demand is small, negative, and not statistically significant. Table 1.5 provides the corresponding mean own and cross-price elasticities of demand for each product, indicating that demand is very inelastic. This finding is consistent with the literature on pharmaceutical demand, closely aligns with the estimates for insulins reported by Einav et al. (2018) for elderly individuals in the US (-0.02), and is not surprising given the French mandated health insurance full coverage of diabetes treatments.

The second part of the table presents the estimated parameters of the normal distributions approximating physicians' prior belief about the preference for new drugs' effect on the average glucose level upon entry, μ_j^0 and V_j^0. The prior means, μ_j^0, are expressed relative to μ_{nj} for the 24-hour insulin product, which is normalized to zero for each patient group. The initial prior means about μ_{nj}, $\widehat{\mu}_j^0$, are significantly negative and lower than their true value, $\widehat{\mu}_{nj}$ (displayed in Table 1.6). These estimates indicate that physicians are initially reluctant to switch existing patients to new products due to a combination of pessimistic priors and potential switching costs.

The estimates display heterogeneity in the mean and variance of physicians' initial beliefs about (new) insulin products. Physicians' beliefs regarding the effectiveness of the 24-hour biosimilar on average glucose levels seem particularly pessimistic and uncertain. In contrast, the 36-hour product exhibits a higher prior mean and a lower variance, which aligns with the characteristics of this new product. This product is indeed derived from the same molecule and produced by the same manufacturer as the 24-hour drug already on the market, for which physicians have accumulated significant experience.[31]

The speed at which physicians learn about the true performance of new drugs on average glucose levels, μ_{nj}, is influenced by the estimated initial prior mean, $\widehat{\mu}_j^0$, the true value, $\widehat{\mu}_{nj}$, the prior variance, $\widehat{V}_j^0$, and the precision of experience signals along with their corresponding variance parameters, $\widehat{\sigma}_0$ and $\widehat{\sigma}_1$. The experience signals are notably imprecise, as their magnitude ranges from 1.6 to 3.8 times that of the initial level of uncertainty. The experience signals received from patients wearing a glucose sensor are equally precise as the experience signals re-

[31] For example, the effects of the existing 24-hour drug on cardiovascular outcomes are well known by physicians who have been prescribing this insulin product since the early 2000s.

ceived without the device as $\widehat{\sigma}_1$ is small and not statistically significant. This empirical finding can suggest that the drug's performance on average glucose levels is accurately approximated by the three-month average glucose level measured in the lab. Figure 1.7 illustrates the evolution of a physician's belief about μ_{nj} on the vertical axis as she accumulates experience signals, represented on the horizontal axis, for different products. The physician's belief is expressed in terms of the deviation from its true value, $\widehat{\mu}^v_{nkj} - \widehat{\mu}_{nj}$, where $\widehat{\mu}^v_{nkj}$ is the empirical mean of the physician's belief after receiving v experience signals. It takes approximately four experience signals for physicians to resolve more than half of the difference between the prior mean and the true value for μ_{nj}.

Table 1.6 and Figure 1.8 display the estimated match values derived from the estimates for μ_{nj}, $\beta_1(x_i)$, and $\beta_2(x_i)$ for three patient groups. These match values reflect the physicians' beliefs about Θ_{ij} for patients with and without a glucose sensor after resolving the initial level of uncertainty regarding μ_{nj}. Within each cluster, the heterogeneity in $\widehat{\mu}_{nj}$ across insulin products illustrates the *perceived* difference in match values without a glucose sensor, while $\widehat{\Theta}_{nj}$ represents the perceived difference with the sensor. By assumption, the perceived difference with the sensor equals the true difference. Table 1.6 presents the estimates of the preference for the drug's effect on the average glucose level, $\widehat{\mu}_{nj}$, for three patient groups. In most instances, $\widehat{\mu}_{nj}$ is significantly greater than zero for new drugs other than the 24-hour biosimilar, indicating that newly introduced products slightly outperform the 24-hour treatment in this regard once uncertainty is resolved.[32] Figure 1.8 displays the physician's match value expectations, $\mathbb{E}(\Theta_{ij})$, with (triangles) and without (squares) a glucose sensor. Figure 1.8 serves as the empirical counterpart to Figure 1.5a. The difference between the average expected match value with and without the sensor corresponds to the average ν_{ij} for the corresponding product j within a group, compared with the (reference) 24-hour product. On average, ν_{ij} accounts for 34% of the clinical match value, Θ_{ij}. The figure highlights that the contribution of the preference for the drug's effect on the glucose profile is heterogeneous across products and patient groups. On average, the difference with the reference good increases for insulin products with a duration of action exceeding 24 hours. The *perceived* match value of these new products improves due to the information generated by the glucose sensor, allowing these products to benefit from the data about the glucose profile provided by the device.

[32]For the 24-hour biosimilar, μ_{nj} is assumed to be equal to the value of the branded version.

Table 1.4: Dynamic parameters

				Mean, μ_j^0		S.D., $(V_j^0)^{0.5}$	
		Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Price (α)		-0.11	(0.20)				
Age (δ)	Cst	-0.52	(0.02)				
	square	0.04	(0.00)				
Prior belief	24-hour biosimilar			-6.05	(0.14)	3.34	(0.11)
	36-hour			-3.43	(0.10)	2.21	(0.05)
	42-hour			-3.49	(0.12)	2.65	(0.08)
	Type 2			-4.25	(0.11)	2.77	(0.09)
Signal S.D.	σ_0	4.30	(0.11)				
	σ_1 (w. CGM)	-0.04	(0.12)				

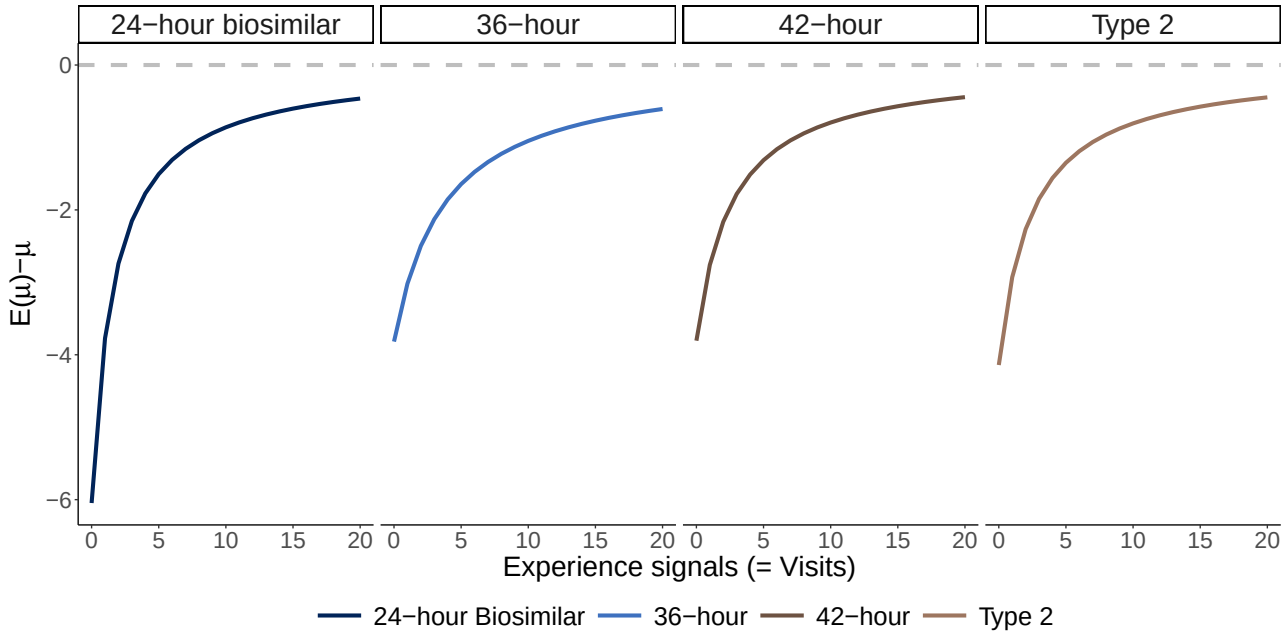
Note: Standard errors are computed from the average of the score. The model is estimated using a sample of 150 diabetes specialists who work outside of hospitals setting.

Table 1.5: Own and cross-price elasticities

	24h	20h	Mix	Human	24h Bios.	36h	42h
Own-price elasticity	-0.050	-0.066	-0.047	-0.046	-0.052	-0.048	-0.049
Cross-price elasticity	0.011	0.003	0.002	0.001	0.002	0.019	0.020

Note: Average across patients.

Figure 1.7: Preference for the drug's effect on average glucose level prior mean, μ_{nj}^{v}



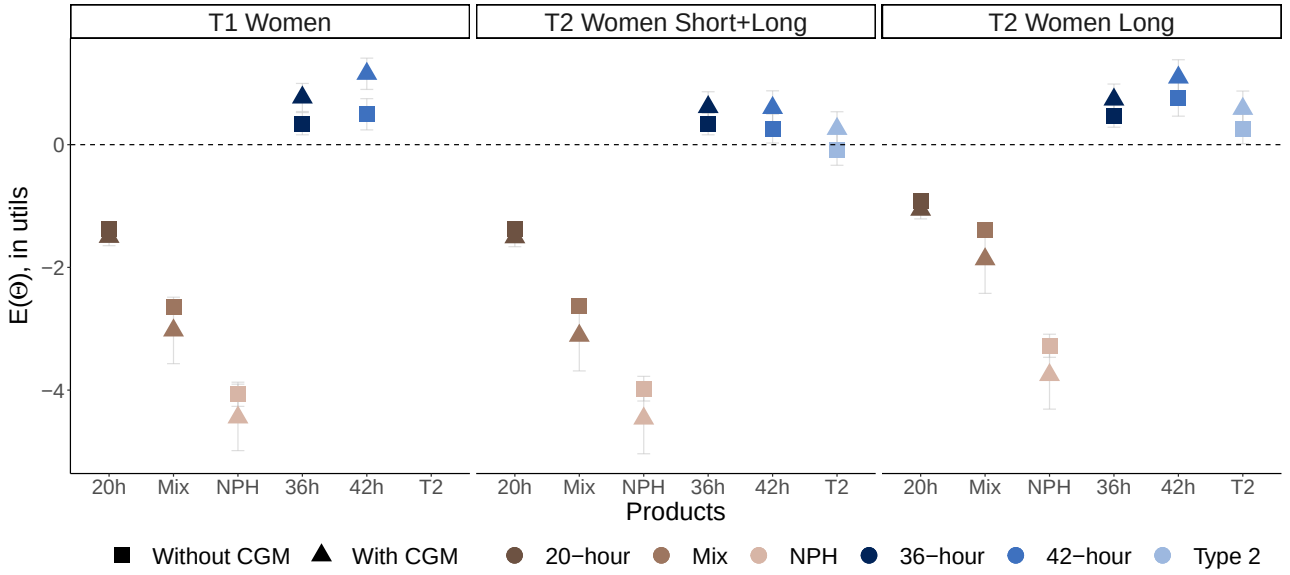
Note: This figure depicts the evolution of a physician's beliefs regarding the preference for the drug's effect on the average glucose level, μ_{nj}, (vertical axis) as the physician gathers more experience signals (horizontal axis). Physicians are Bayesian and update their (normally distributed) prior beliefs based on signals received when a patient returns to her practice while using the new product j. The parameters of the prior and signal distributions are estimates derived from the demand model (Table 1.4). The figure illustrates the beliefs for Type 2 male patients undergoing short- and long-acting insulin therapy, assuming that signals are collected from patients not utilizing a CGM. The vertical axis represents the difference between the prior mean at v, $\widehat{\mu}_{nj}^{v}$, and the true value, $\widehat{\mu}_{nj}$, which converges to zero as physicians accumulate more experience.

Table 1.6: Preference for the drug's effect on average glucose level, μ_{nj}

		Old drugs			New drugs		
		20-hour	Mix	Human	36-hour	42-hour	Type 2
Type 1 'Old male'	Coef	-1.72	-2.43	-4.37	0.31	0.40	
	S.E.	0.03	0.06	0.12	0.11	0.15	
Type 2 short+long 'Male'	Coef	-1.53	-2.53	-4.41	0.39	0.31	-0.10
	S.E.	0.03	0.06	0.13	0.10	0.14	0.15
Type 2 long 'Male'	Coef	-1.12	-1.52	-3.31	0.03	0.47	0.23
	S.E.	0.02	0.06	0.11	0.11	0.17	0.13

Note: Standard errors are computed from the average of the score. Type 1 and Type 2 refer to the patient's type of diabetes. Type 1 patients must use both short-acting and long-acting insulins daily. Type 2 patients may depend on either short-acting and long-acting insulins (referred to as 'short+long') or long-acting insulin only ('long').

Figure 1.8: Perceived match value with/without a CGM



Note: This figure presents the perceived match value, $E(\Theta_{ij})$, both without the technology (■) and with a CGM (▲) in utils (vertical axis) for each product (horizontal axis). The expectation assumes that $\mathbb{E}(\mu_{ij}) = \mu_{nj}$ for patient i in group n, resolving the initial uncertainty surrounding μ_{ij} for new drugs. The perceived match value without the technology corresponds to μ_{nj}. In contrast, the perceived match value with the technology is $\mu_{nj} + \nu_{nj}$, where ν_{nj} is the average across patients within a cluster. The 24-hour product (not shown) corresponds to the normalized good in each group. 'Mix' refers to insulin mixes, 'NPH' denotes human insulins, and 'Type 2' designates the combination of long-acting insulin and another molecule, meant exclusively for type 2 diabetes patients. New drugs are represented in blue, while old products are in brown. Only three clusters out of seven patient groups are displayed. The remaining clusters are presented in Figure 1.32.

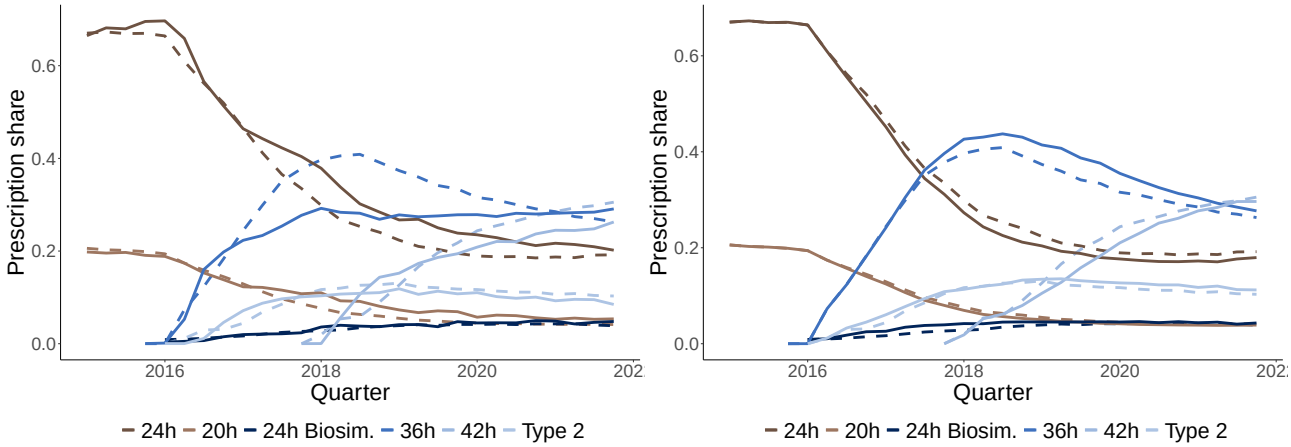
Model fit

Before presenting the pricing model and counterfactual analysis, this section provides descriptive evidence of how well the demand model estimates align with the data. The sets of medical appointments assigned to a particular physician k, the identity of the patient i, and the choice of glucose monitoring technology, a_{iv}, are kept constant. I use the parameter estimates to simulate the choices and choice probabilities for each product at every medical appointment v. When comparing the model's predictions to the actual data, I fix the previous insulin product choices, $j_{i,v-1}$, to update physicians' prior beliefs. This assumption prevents the accumulation of prediction errors over time. I also compute the average predicted choice probabilities in a 'fully simulated' environment, where the predicted choice at v influences future prescription learning, to illustrate the empirical model's capability to reproduce the diffusion pattern observed in the data.

Figure 1.9 displays the average simulated choice probabilities per product per quarter, along with the empirical insulin product shares in Figure 1.9a, and the 'fully simulated' average choice probabilities in Figure 1.9b. The parameters generally overestimate prescriptions for the 36-

hour product compared to the actual choices made by physicians during the years immediately following the product's introduction. To a lesser degree, the model slightly overestimates prescriptions for the 42-hour drug from 2019 to 2021. Figure 1.33 compares realized and predicted choices, while Table 1.16 shows the frequency of accurate predictions across years and patient groups. The accuracy rates range from 31% to 82%, similar to those reported in Dickstein (2021), who estimates a demand model with learning for antidepressants in the US.

Figure 1.9: Model fit for product shares



(a) Realized and predicted (dashed) shares (b) Fully simulated and predicted (dashed) shares

Note: Figure 1.9 compares each product's actual and predicted choice probabilities over time. For clarity, insulin mixes and human insulins are not represented.

1.5 Drug price setting model

This section presents and estimates the primitives of the insulin price setting model. This framework considers the pricing responses from drug manufacturers and the regulator following the introduction of a digital medical device. As noted in the previous section, the arrival of innovations in drugs and devices is taken as given, motivated by the timeline of drug and device development.[33]

1.5.1 Empirical model

Even when decision-makers are myopic, as assumed in section 1.4, a demand system subject to consumer learning presents dynamic features. Current sales influence future demand. Model-

[33]Drug development can take more than 10 years. One of the first patents for Abbott Freestyle Libre approved by the EMA in 2014 dates back to 2006/2007 (Litvinova et al. (2023)).

ing supply in markets with dynamic demand is inherently complex, as firms have incentives to leverage this feature and set their prices in a forward-looking manner (Shapiro (1983), Bergemann and Välimäki (2006)). A firm's optimal pricing strategy relies on the distribution of the perceived match values of both its own and competing products among consumers. To overcome the complexities posed by forward-looking drug manufacturers and dynamic demand, I rely on the characteristics of the institutional setting. In France, the price for prescription drugs is determined through negotiations between drug manufacturers and the regulator. The price is established upon entry for an extended period and is renegotiated over time. Price increases are difficult to negotiate, preventing manufacturers from taking advantage of the market dynamics by setting a low price upon introduction to raise it in later periods in the spirit of Shapiro (1983).

I assume that the regulation results in an equilibrium price as the outcome of a multilateral Nash bargaining process between drug manufacturers and the regulator, similar to Tunçel (2024). The Nash-in-Nash bargaining solution considers the negotiation outcome between other drug manufacturers and the regulator as given. In pharmaceutical markets, each branded insulin product is produced by a single manufacturer. I treat the 24-hour biosimilar and its branded version as separate products.[34] Unlike other pharmaceuticals, the insulin market is concentrated, with three companies providing the complete range of products in France. Bargaining takes place at the drug portfolio level for each manufacturer. The profit for firm f offering products $j \in \mathcal{J}_{ft}$ in year t is

$$\pi_{ft}(\mathbf{p}_t) = \sum_{\forall j \in \mathcal{J}_{ft}} (p_{jt} - c_{jt}) q_{jt}(\mathbf{p}_t) \tag{1.17}$$

where $q_{jt}(\mathbf{p}_t) = \sum_{\forall i,k} \sum_{\forall v \in t} s_{ikjv}(\mathbf{p}_t)$ is the total demand for drug j across appointments v occurring in year t, as estimated in section 1.4. This profit function assumes that manufacturers have perfect foresight about demand realization in the next 12 months following the price negotiation. The manufacturer anticipates that physicians' learning alters the initial market shares for new products upon entry. Individual choice probabilities are derived from Equation 1.13. From 2017 onward, drug manufacturers also have rational expectations about the CGM insurance coverage decision, the adoption among the pool of participants, a_{iv}, and how the digital device shifts demand via ν_{ij}. This last assumption presupposes that, when bargaining

[34] These products are not 'interchangeable' by pharmacists in France during my sample period. The physician writes a different prescription for each molecule version, and the pharmacist must provide the product specified on the prescription.

over prices, drug manufacturers have superior information to that of physicians about the clinical benefits of each own and competing new drug in real-life conditions.[35] Pharmaceutical companies know their marginal costs, c_{jt}, when negotiating prices in year t.

When bargaining with firm f, the regulator is assumed to maximize the *ex-ante* consumer surplus generated by the portfolio of drugs offered by firm f in year t, given by

$$\Delta_{ft}CS(\mathbf{p}_t) = \sum_{\forall i,k}\sum_{\forall v\in t}\frac{1}{\lambda}\ln\Big(\sum_{\forall j\in\mathcal{J}}\exp(u_{ikjv}(p_{jt(v)},d_j,a_{iv}|\mathcal{I}_{kv})\Big) - \sum_{\forall i,k}\sum_{\forall v\in t}\frac{1}{\lambda}\ln\Big(\sum_{\forall j'\notin\mathcal{J}_f}\exp(u_{ikj'v}(p_{j't(v)},d_{j'},a_{iv}|\mathcal{I}_{kv})\Big) \tag{1.18}$$

The parameter λ corresponds to the scaling factor for consumer surplus from the regulator's perspective. The scaling factor, λ, can differ from α, allowing the price sensitivity of physicians, estimated from the demand model, not to accurately convert utils into euros for the regulator. The regulator does not internalize the impact of pharmaceutical prices on innovation through profits. Under Nash bargaining, both parties have symmetric information, meaning that the regulator knows the firm's marginal costs. The regulator also forms rational expectations about the adoption of the digital device and the impact of the device on consumer surplus when bargaining over insulin prices from 2017 onwards. I follow Grennan and Town (2020) in assuming that the surplus the regulator considers when setting drug prices does not account for the difference between the decision and experience utility. Under these assumptions, access to the digital device affects equilibrium prices through profits and the consumer surplus generated by the reimbursement of certain products.[36] Denoting b_{ft} the bargaining ability of firm f in year t, the Nash-in-Nash equilibrium prices maximize the Nash product for the manufacturer's profit and the regulator surplus, taking the prices of other products as follows:

$$\max_{\mathbf{p}_{jt},j\in\mathcal{J}_{ft}}[\pi_{ft}(\mathbf{p}_t)]^{b_{ft}}[\Delta_{ft}CS(\mathbf{p}_t)]^{1-b_{ft}} \tag{1.19}$$

where the disagreement profits for the pharmaceutical company is zero. The portfolio of each pharmaceutical company is treated as an indivisible block, exogenously given. I do not consider bargaining over a subset of products. The first-order condition with respect to the price of drug

[35] Hitsch (2006) and Handel and Misra (2015) study firms' pricing decisions when they have incomplete information about the true demand curve but assume that no learning occurs on the demand side. Firms are assumed to have accurate information about the clinical match value of new products in real-life conditions.

[36] In France, the regulator rewards pharmaceutical products for enhancing patient outcomes compared with existing alternatives through higher prices.

$j \in \mathcal{J}_{ft}$ in year t is given by

$$b_{ft}\frac{\partial \pi_{ft}(\mathbf{p}_t)/\partial p_{jt}}{\pi_{ft}(\mathbf{p}_t)} + (1-b_{ft})\frac{\partial \Delta_{ft} CS(\mathbf{p}_t)/\partial p_{jt}}{\Delta_{ft} CS(\mathbf{p}_t)} = 0 \tag{1.20}$$

The scaling factor of the consumer surplus, λ, does not affect the equilibrium outcome, whereas the price sensitivity of demand matters. When a firm offers two products, j and j', the first-order conditions yield the following pricing equation:

$$p_{jt} = c_{jt} - \left[\beta_{ft} h_{jt} + \frac{\partial q_{jt}(\mathbf{p}_t)/\partial p_{jt}}{q_{jt}(\mathbf{p}_t)} + \left(\beta_{ft} h_{j't} + \frac{\partial q_{j't}(\mathbf{p}_t)/\partial p_{jt}}{q_{jt}(\mathbf{p}_t)}\right) \frac{\left(q_{j't}(\mathbf{p}_t)\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}} - q_{jt}(\mathbf{p}_t)\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}}\right)}{\left(q_{jt}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} - q_{j't}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}}\right)}\right]^{-1} \tag{1.21}$$

where $h_{jt} = \frac{\partial \Delta_{ft} CS(\mathbf{p}_t)/\partial p_{jt}}{\Delta_{ft} CS(\mathbf{p}_t)} < 0$ and $\beta_{ft} = \frac{1-b_{ft}}{b_{ft}}$. For single-product firms,

$$p_{jt} = c_{jt} - \left[\beta_{ft} h_{jt} + \frac{\partial q_{jt}(\mathbf{p}_t)/\partial p_{jt}}{q_{jt}(\mathbf{p}_t)}\right]^{-1} \tag{1.22}$$

The details for the first-order condition computation are provided in Appendix 1.11. The unknowns from Equation (1.21) are the marginal costs, c_{jt}, and the bargaining parameters, b_{ft}, as the remaining elements are observed or can be computed from the estimates of the demand system. These unknown primitives must be recovered to compute the equilibrium insulin prices under alternative scenarios. Yet, the first-order conditions generate a system of $\mathcal{J}_{ft}$ equations and $\mathcal{J}_{ft}+1$ unknown, leading to an identification challenge common in the Nash bargaining literature (Grennan (2013), Tunçel (2024)).

1.5.2 Estimation and results

To overcome this identification challenge and recover the primitives of the supply model, it is further assumed that

$$\begin{aligned} c_{jt} &= \gamma mc_j + \zeta_{jt} \\ \frac{1-b_{ft}}{b_{ft}} &= \beta_f \end{aligned} \tag{1.23}$$

mc_j are molecule-level costs of production for a daily dose in 2016, using the estimates from Gotham et al. (2018), and ζ_{jt} is an unobserved cost shock.[37] The bargaining weights are assumed to be firm-specific and constant over time. By combining Equation (1.21) and the restrictions from Equation (1.23), we obtain the following:

$$\zeta_{jt} = p_{jt} - \gamma mc_j + \left[\beta_f h_{jt} + \frac{\partial q_{jt}(\mathbf{p}_t)/\partial p_{jt}}{q_{jt}(\mathbf{p}_t)} + \left(\beta_f h_{j't} + \frac{\partial q_{j't}(\mathbf{p}_t)/\partial p_{jt}}{q_{jt}(\mathbf{p}_t)}\right) \frac{\left(q_{j't}(\mathbf{p}_t)\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}} - q_{jt}(\mathbf{p}_t)\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}}\right)}{\left(q_{jt}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} - q_{j't}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}}\right)}\right]^{-1} \tag{1.24}$$

The above equation underscores the common endogeneity concern that arises on the supply side between the price and the marginal cost shock. I rely on traditional instruments that are correlated with price but uncorrelated with the idiosyncratic cost shock, such as the number of competing products. By including firm fixed effects and marginal costs in the set of exogenous variables, I can estimate the bargaining weights and marginal cost parameter via GMM using the following moment conditions:

$$E[\zeta_{jt}|Z_{jt}] = 0 \tag{1.25}$$

In the main specification, the coefficient for production costs, γ, is set to 1.3, as it is difficult to estimate precisely both marginal costs and firm-specific bargaining weights. To set γ, the model is estimated for $\gamma \in [1, 2]$, and the value minimizing the GMM objective function is kept in the main specification. An alternative specification could assume that $\gamma = 1$; however, several reasons rationalize $\gamma > 1$. I rely on the 'competitive' estimates from Gotham et al. (2018). This can also be explained by the cost of the injection device, which is not included in the current estimate, and distribution costs.[38] The results are shown in Table 1.7. Insulin mixes (including the new Type II drug) and human insulins are excluded from the estimation, keeping their prices fixed in the counterfactuals.

The bargaining weights reflect the drug manufacturer's capacity to set a price above its marginal cost when negotiating with the regulator. The estimated manufacturer-level bargaining weights range from 5 to 7%. These figures are significantly lower than those estimated for various prescription drugs by Dubois et al. (2022) in Canada and Tunçel (2024) in France regarding the antidepressant market. This discrepancy may be driven by differences in the price

[37] Gotham et al. (2018) estimates the cost of production for a vial containing 1,000 insulin units relying on Indian customs data for raw molecules and excipients' quantities and prices. Among others, the marginal cost to consider in the demand model may deviate from these estimates as pens, the most common injection device, are more costly than vials.

[38] Most patients in France rely on insulin pens to inject their insulin, while the current estimates are calculated for vials. Robustness checks can be conducted to account for the cost of the injection device.

elasticity of demand, and the specificity of diabetes insurance coverage. With demand being nearly inelastic to prices, the regulator serves as the primary force preventing pharmaceutical companies from setting excessively high prices.[39] The average per-product margin implied by the model lies between 53 and 85%.

Table 1.7: Bargaining weights estimates

	Coef.	S.E.	b_f
Firm 1	15.383	0.914	0.061
Firm 2	19.031	2.722	0.050
Firm 3	13.360	1.811	0.070
γ		1.300	
N		30	

Note: One-step GMM. Jackknife standard errors. γ is set to 1.3 by comparing the value of the GMM objective function for $\gamma \in [1, 2]$. The estimation excludes insulin mixes, human insulin and the Type 2 product, whose prices are held fixed in counterfactuals.

1.6 Data-driven insights, physician learning and cross-market complementarities

In this section, the estimates from the demand and supply model are used to assess the impact of the insights generated by digital wearables on pharmaceutical demand through various counterfactual scenarios. The equilibrium framework aims to understand: (i) the effect of information provision on insulin product demand and pricing; (ii) the losses incurred from relying on partial information when glucose sensors are unavailable; and (iii) how CGMs information affects the profitability of pharmaceutical innovations. In each counterfactual scenario, I compute the new equilibrium drug prices using the price setting model and the realization of demand.

1.6.1 Defining relevant market outcomes

Before implementing and comparing counterfactual scenarios, I define the relevant indicators to understand whether and how the market dynamic shifted owing to the digital device. This section focuses on four key indicators of market outcomes: (i) market shares, (ii) firms' profit, (iii) physician-level learning, and (iv) consumer welfare.

The first two market outcomes are straightforward to compute from the predicted choice

[39] In countries such as the US, where the government does not intervene in drug pricing, insulin prices for the same product offered in France are significantly higher.

probabilities and equilibrium prices. Physician-level learning is studied through physicians' end-of-period beliefs about μ_{nj}, denoted $\mu_{nkj}^{V_k}$. Specifically, I compute the difference between the estimated preference for drug j's effect on the average glucose level for patient group n, $\widehat{\mu}_{nj}$, and the average of physician k's belief about μ_{nj} in the last appointment, V_k, denoted $\widehat{\mu}_{nkj}^{V_k}$. This difference is a proxy for the accuracy of physicians' beliefs at the end of the sample period. In the model, as the physician accumulates experience signals, the average of her belief approaches the true value. Therefore, a smaller difference indicates a more 'accurate' belief about product j.

Computing consumer welfare under different scenarios is more intricate, as decision-makers face imperfect information about the patient-product clinical benefits of each drug. As mentioned in section 1.4, a discrepancy exists between the decision and experienced utility. Denoting $\mathbf{p}$ the equilibrium prices, $\mathbf{d}$ the vector of product durations, d_j, and $\mathbf{a}$ the vector of CGM usage in the patient population, the expected indirect utility for patient i from the choice made on his behalf by physician k at time v is as follows:

$$\begin{aligned} W_{ikv}(\mathbf{p},\mathbf{d},\mathbf{a}) &= \mathbb{E}_{\varepsilon}[\bar{u}_{ikj^*v}(\mathbf{p},\mathbf{d},\mathbf{a})] \\ &= \ln\Big(\sum_{\forall j}\exp(u_{ikjv}(p_{jt(v)},d_j,a_{iv}|\mathcal{I}_{kv}))\Big) - \sum_{\forall j} s_{ikjv}(\mathbf{p},\mathbf{d},\mathbf{a})\Big[\mu_{nkj}^{v} - \mu_{nj} + \nu_{ij}(d_j)(a_{iv}-1)\Big] \end{aligned} \tag{1.26}$$

where $s_{ikjv}(\mathbf{p},\mathbf{d},\mathbf{a})$ denotes the choice probability, given by Equation (1.13). The value of learning is taken into account in the second component of the indirect utility, which accounts for the difference between the expected match value at the time of the decision, $\mu_{nkj}^{v} + \nu_{ij}a_{iv}$ and its true value, $\Theta_{ij} = \mu_{nj} + \nu_{ij}$. This expression allows for a comparison of consumer welfare under alternative market outcomes, m, from the compensating variation.[40] The welfare induced by the digital device for the patients is beyond what is measured in this project, as I restrict my attention to the impact of the device on the long-acting insulin market.

[40]The compensating variation is computed as

$$CV_{ikv}(\mathbf{p}^m,\mathbf{d}^m,\mathbf{a}^m) = \frac{1}{\alpha}\big(W_{ikv}(\mathbf{p}^m,\mathbf{d}^m,\mathbf{a}^m) - W_{ikv}(\mathbf{p}^0,\mathbf{d}^0,\mathbf{a}^0)\big)$$

where $W_{ikv}(\mathbf{p}^m,\mathbf{d}^m,\mathbf{a}^m)$ correspond to consumer welfare in scenario m and $W_{ikv}(\mathbf{p}^0,\mathbf{d}^0,\mathbf{a}^0)$ correspond to consumer welfare without the technology.

1.6.2 Digital device adoption and the insulin market

The model is used to evaluate the impact of CGM on the insulin market, taking patient adoption of the digital device as given. To that extent, I simulate a scenario where CGMs are no longer used, setting a_{iv} to zero for all patients. I compare the market equilibrium with and without CGM adoption.

Figure 1.10 illustrates the difference in consumer welfare induced by the introduction of CGMs computed from Equation 1.26. Figure 1.10a shows the compensating variation (vertical axis) in euros per day following a prescription occurring in period t (horizontal axis). The average consumer welfare is presented for patients with a CGM (blue curve) and those not using a CGM (brown curve). The average is computed by including all patients (solid lines) and by restricting to patients eligible for CGM insurance coverage (dashed lines). Three key insights emerge. First, the average welfare gains for CGM users are nearly ten times greater than those for nonusers, indicating limited information spillovers between the two groups. Second, among CGM users, welfare gains are greater in the early months of CGM coverage and decline over time. The information generated by the technology is most valuable for patients shortly after drugs' entry when physicians face greater uncertainty about the performance of new products on real-life patients. Despite the decreasing trend, CGMs should continue generating positive consumer welfare gains for users in the long run — as physicians gather sufficient experience about new drugs — due to the information about the glucose profile, ν_{ij}. Third, for nonusers, the welfare gains take longer to materialize. Physicians must observe the performance of new drugs from patients with the device for the information to benefit nonusers. Over time, nonuser gains should converge to zero as physicians accumulate enough knowledge about new drugs. Figure 1.10b displays the distribution of consumer welfare gains across patients using CGM, distinguishing between individuals with Type I and Type II diabetes and with different demographic characteristics. Patients with Type I diabetes benefit more than those with Type II diabetes do, and women tend to benefit more than men within each diabetes type. These findings align with the medical literature, which suggests that patients with a long history of diabetes and women are more prone to low glucose levels overnight, a condition for which CGMs provide relevant insights (Siamashvili et al. (2021)) and triggered by some insulin products already on the market.

Next, I assess the impact on physician learning. Figure 1.11 compares the difference between $\hat{\mu}_{nj}$ and $\hat{\mu}_{nkj}^{V_k}$, the preference for the drug's effect on the average glucose level and the end-of-

period physician-level prior mean. This difference is computed for each physician, k, patient group, n, and new product j. The figure plots belief 'accuracy' without the device (horizontal axis) against its value when the technology is available (vertical axis).[41] Points above (below) the 45-degree line indicate more (less) accurate beliefs about product j when the device is available. Interestingly, while CGMs accelerated learning about some products (e.g., the 42-hour product), they did not enhance learning for others (e.g., the 24-hour biosimilar). Three factors drive this heterogeneous impact: (i) the contribution of ν_{ij} to the match value, Θ_{ij}, is heterogeneous across insulin products and new drugs with a duration of action exceeding 24 hours have a higher perceived match value owing to the information generated by glucose sensors; (ii) physicians face a limited number of opportunities to gain experience with new drugs such that products compete over these opportunities; and (iii) the experience signals from CGM users about μ_{nj} are as precise as those received from regular patients. By generating new observable drug attributes, such as a drug's performance regarding overnight glucose levels, for both old and new products, glucose sensors direct insulin demand for CGM users toward drugs that excel in these dimensions. These products are also new to the market. As patients return to their physician while using the new product, the physician learns about μ_{nj} from their experience. This opportunity to learn about μ_{nj} would not have arisen if the patient had not been switched to that drug based on the information provided by his sensor (ν_{ij}). However, because the number of appointments fixes the number of learning opportunities and experience signals from CGM users are not more precise than those from patients without the device, the physician learns slightly more slowly about new drugs that do not excel in these new observable attributes. Consequently, learning about some products may come at the expense of others, and not all new products benefit equally from CGMs' insights.

At the market level, Table 1.8 shows the market shares for 2021, whereas Figure 1.12 tracks their evolution over time. Figure 1.13 presents the manufacturers' profits. In 2021, the 42-hour product experienced a 16.5% increase in market share due to CGM adoption. In contrast, the 36-hour product experienced only a temporary increase in market share, which nearly vanished by the end of 2021. Similarly, profits for the 42-hour product rose by 23%, driven primarily by the demand response to CGMs (+18%). These results suggest that CGMs unevenly favor new molecule adoption at the expense of older molecules, particularly 24-hour products.

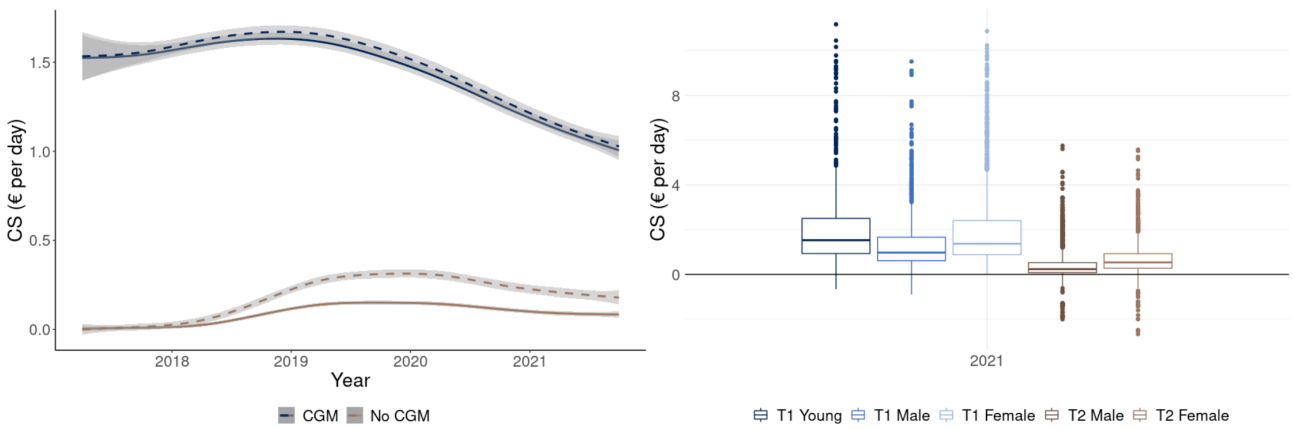
In summary, this counterfactual analysis shows that the introduction of CGMs (i) primarily

[41] In this context, prior beliefs are inaccurate relative to the true partial information match value. The results from the estimation suggest pessimistic prior beliefs since $\widehat{\mu}_j^0 < \widehat{\mu}_{nj}$ for each new drug j and patient group n.

benefited CGM users with limited spillovers to nonusers, (ii) accelerated physician learning about specific products, and (iii) was sufficient to affect market shares and insulin manufacturers' profits.

The limited spillovers to nonusers may be influenced by the unequal propensity of physicians to see CGM patients (Table 1.1).[42] To explore this further, I conduct a counterfactual reallocation of glucose sensors among eligible patients, accounting for patient demographics that may correlate with adoption (details in Appendix 1.12). Figure 1.34 shows that reallocating sensors does not increase spillovers to nonusers. This finding indicates that the variation in CGM adoption across physicians is not the primary driver of limited spillovers. However, this counterfactual does not address how spillovers could be enhanced by allocating sensors to patients with *different* demographic characteristics.

Figure 1.10: Consumer welfare



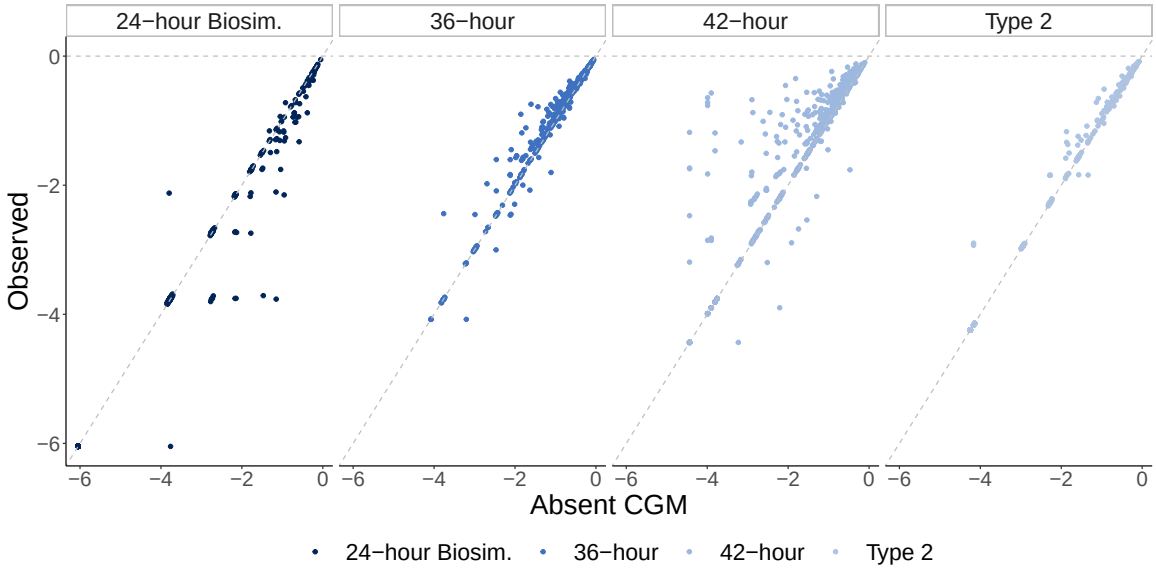
(a) Average over time

(b) Distribution across patient types

Note: Figure 1.10a presents the compensating variation (vertical axis), in euros per day, following a prescription occurring in period t (horizontal axis). The average consumer welfare is presented for patients with CGMs (blue curve) and without CGMs (brown curve). The dashed lines focus on the consumers eligible for the device. Figure 1.10b presents the compensating variation distribution (vertical axis) across different eligible patient types (horizontal axis) in 2021 for CGM users. Ineligible clusters are not represented here as the effect is negligible.

[42] In the extreme case where all patients in the same cluster are either adopting or not at a physician practice, spillovers cannot arise.

Figure 1.11: Learning about the drug's effect on the average glucose level, $\hat{\mu}_{nj}$, with and without CGMs



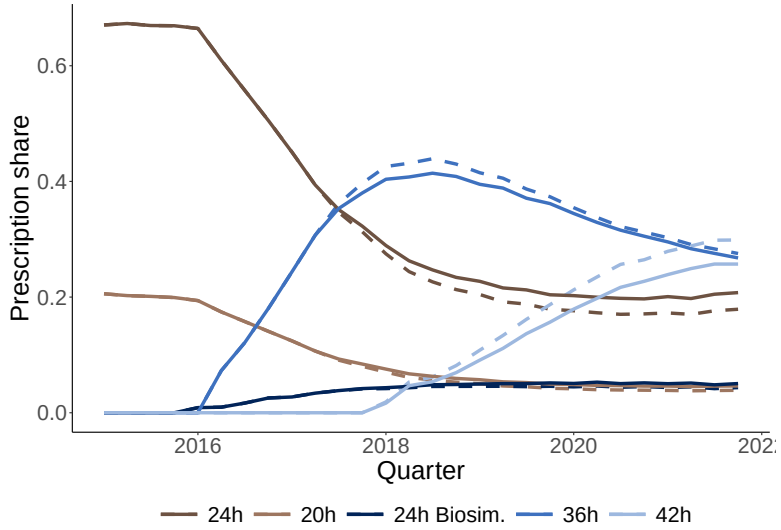
Note: This figure plots the difference between $\hat{\mu}_{nj}$ and $\hat{\mu}^{V_k}_{nkj}$, the partial information clinical match value and the end-of-period physician-level prior mean in an environment with CGMs (vertical axis) against its value in an environment without CGMs (horizontal axis). One observation per physician, patient type among types eligible for the technology and product. Points above the 45-degree lines suggest more accurate belief at the end of the period when the technology is available.

Table 1.8: Market shares with vs without CGM in 2021

Scenario	24-hour	20-hour	24-hour Bios.	36-hour	42-hout
Base (no CGM)	0.203	0.046	0.050	0.281	0.250
Demand-response	0.174	0.038	0.044	0.287	0.292
+ Supply response	0.174	0.038	0.043	0.289	0.291

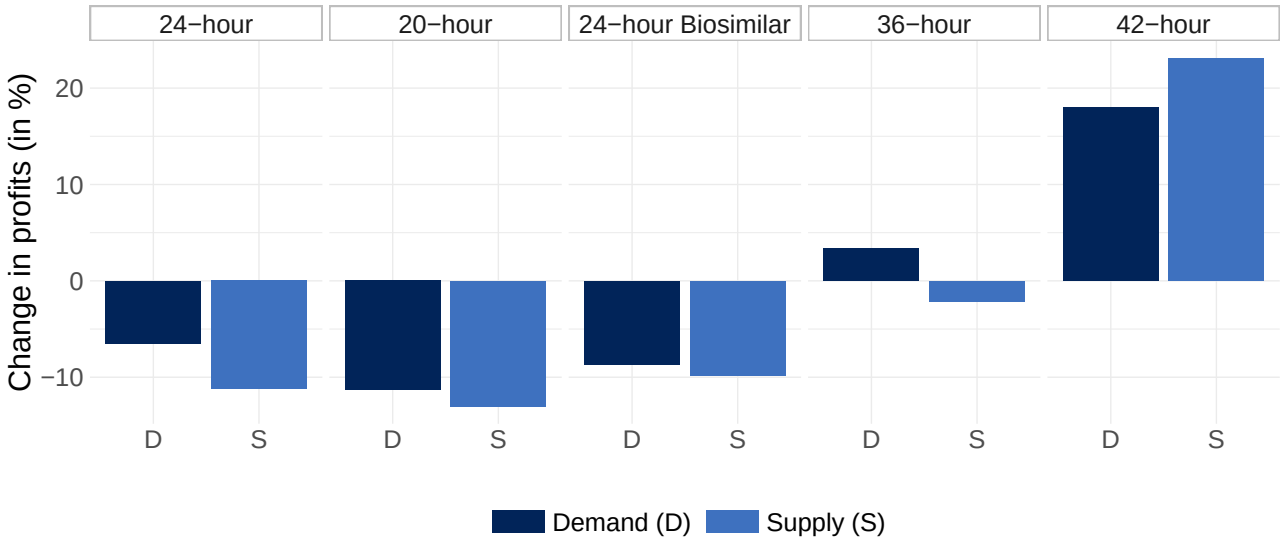
Note: The first row corresponds to the predicted 2021 market share, absent the technology. The second row takes into account the demand response to CGM information. The third row allows prices to respond to changes in demand and correspond to the predicted 2021 market shares, given the observed adoption pattern. Insulin mixes (including Type II product) and human insulin are not represented.

Figure 1.12: Market shares without vs with (dashed) CGMs over time



Note: This figure plots the choice probabilities without CGMs (plain lines) and their counterparts as CGMs became available (dashed lines) over time. Insulin mixes (including the new Type II product) and human insulin shares are not represented.

Figure 1.13: Impact of CGMs on drug manufacturers' profits from 2017-2021



Note: Change in profits due to the introduction of CGMs from 2017-2021 for each product. The profits without CGMs are normalized to 100 for each product, and each scenario is compared to this baseline. For each product, the first bar corresponds to changes in profits due to CGMs affecting the demand curve, and the second bar accounts for prices to react to the change in demand.

1.6.3 Losses from relying on partial information about the clinical match value

By generating continuous glucose level data, the digital device emphasizes the inefficiencies arising from incomplete learning in prescription drug choice. Observing the choice for patients under complete learning helps quantify the costs of relying solely on partial information, in terms of market shares, foregone profits and consumer welfare. In this section, I use the estimates for ν_{ij} to measure the losses from incomplete learning. I assess how much of this information gap is bridged by providing broad access to CGMs.

I begin by considering the complete information scenario in which every patient uses a CGM as the 'frictionless' solution. In this case, physicians access each patient's true clinical match values, Θ_{ij}. I then compare this market outcome to two scenarios where (i) the technology was never available and (ii) CGM coverage is available, but some patients do not adopt the device, reflecting the current situation. Comparing the frictionless case to the first scenario, in which choices rely exclusively on partial information, emphasizes the losses from incomplete learning in terms of industry profits and consumer welfare. To abstract away from the uncertainty about new drugs' performance upon market entry, I focus on the steady-state choice probabilities that prevail once this friction is resolved. The steady-state choice probabilities are obtained by assuming physicians receive a precise signal from patients about the performance of new drugs ($\sigma_0 = 0.001$ and $\sigma_1 = 0$) and restrict my attention to the last year of the data, over 36 months after the last product entered the market.[43] The new equilibrium prices are computed for each drug under each scenario. Table 1.9 shows each product's steady-state shares, and Figure 1.14 presents the profits. A comparison of the first two rows highlights the products winning (older drugs) and those losing out (new products) due to incomplete learning. When these losses are measured in terms of consumer welfare, partial information accounts for -0.36€/per day on average across prescriptions — 8 times less than the daily cost of continuous glucose monitoring. Providing broad CGM coverage for patients closes 54.8% of this gap, slightly above the adoption rate, since patients using CGMs experience greater losses from partial information (Figure 1.15).

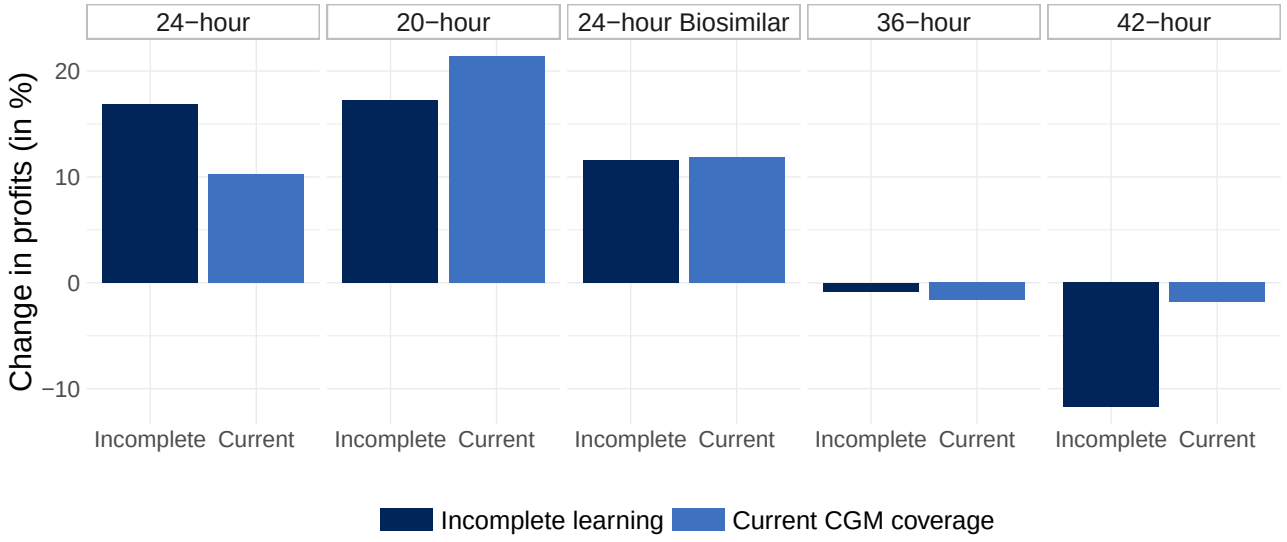
[43] Given the initial prior variance and the signal precision presented in Table 1.4, the uncertainty level becomes negligible after a physician receives the first feedback.

Table 1.9: Steady state market shares in 2021, complete vs partial information

	Product				
Complete			New	New	New
Information	24-hour	20-hour	24-h biosimilar	36-hour	42-hour
Yes	0.095	0.020	0.076	0.271	0.383
No	0.122	0.028	0.095	0.264	0.332
Users	0.106	0.024	0.085	0.267	0.363

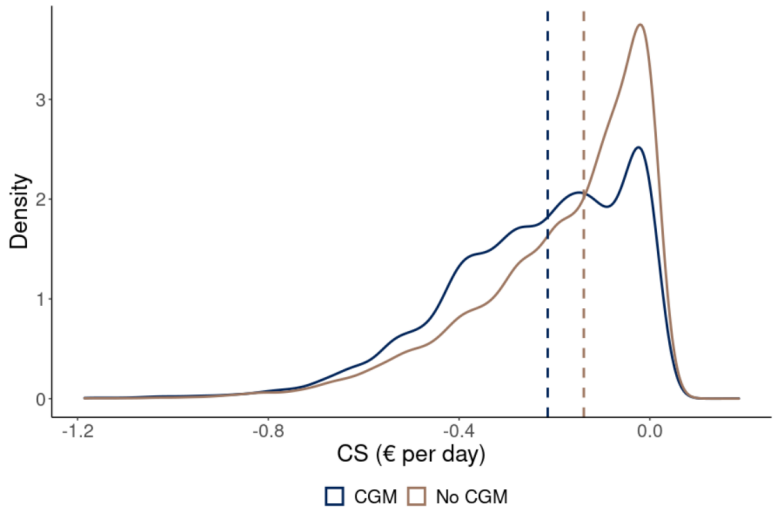
Note: This table presents market shares for each product once the uncertainty about the value of θ_{nj} for new products is resolved. Each row differs by the set of patients using CGMs daily. Insulin mix (including the Type 2 product), human insulin are not represented as their price are held fixed in counterfactuals.

Figure 1.14: Changes in profits due to incomplete learning



Note: Change in profits for 2021 compared with the 'frictionless' case where the choice is made under complete learning for each patient. Each bar corresponds to the profit from product j in 2021 in a given scenario. 'Incomplete learning' corresponds to the scenario in which glucose sensors are not available. 'Current CGM coverage' corresponds to the scenario in which glucose sensors insurance coverage is available and some patients adopt, reflecting the current situation. The profits in the 'frictionless' case are normalized to 100 for each product, and each scenario is compared to this baseline. Bars above (below) zero suggest higher (lower) profits than if the choice was made under complete learning for all patients. For each product, the first bar corresponds to changes in profits in the absence of CGMs, and the second bar corresponds to the changes in profits under the current coverage and adoption decisions.

Figure 1.15: Losses from partial information



Note: The figure plots the distribution of compensating variations between the frictionless case and relying exclusively on partial information across patients. Distributions are plotted separately for patients who adopted the technology in the observed coverage scheme (blue curve) and those who did not (brown curve). The top and bottom 2.5% are not represented here for clarity. The vertical dashed line corresponds to the median loss for each patient population.

1.6.4 Cross-market complementarities

The results from the previous sections indicate that the introduction of CGMs did not benefit all the drugs equally. In 2021, the market share of the 42-hour product was 16.2% higher owing to the insights generated by CGMs, compared with only 3% for the 36-hour product.

This result suggests that CGMs better leverage the potential of certain new drugs. This section documents how innovation in medical devices, which provide new observable drug attributes, impacts the value of new products in pharmaceutical markets. The aim of this analysis is to quantify the extent of benefits derived from the introduction of new drugs and medical devices due to the complementarity between the data insights generated by the device and the characteristics of insulin products. Following Petrin (2002), I assess the benefits of new products using consumer welfare.

Specifically, I simulate market outcomes assuming that each new product has not entered and CGMs are unavailable while keeping competitors' price and entry decisions unchanged. From the baseline scenarios in which one new product and CGMs are not available, I consider unilateral deviations in the entry decision for the 36-hour and 42-hour products and in the availability of CGMs for diabetic patients.[44] For each product j and patient i, I denote $\Delta CS_{ij}^{e,a}$

[44]The price of competing products is fixed to abstract away from the consumer welfare induced by the

as the change in consumer welfare in situations where the product enters or not, $e \in \{1, 0\}$, and whether or not the technology is available, $a \in \{1, 0\}$, compared with the baseline scenario in which the drug did not enter ($e = 0$) and CGMs are unavailable ($a = 0$). When $a = 1$, the adoption of glucose sensors in the patient population corresponds to the one observed in the data. I consider the following decomposition:

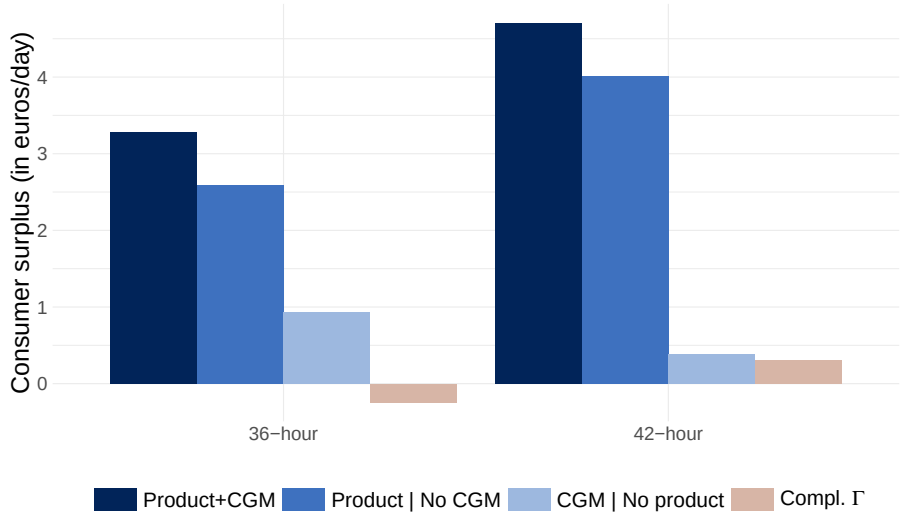
$$\Delta CS_{ij}^{1,1} = \Delta CS_{ij}^{1,0} + \Delta CS_{ij}^{0,1} + \Gamma_{ij} \tag{1.27}$$

where $\Delta CS_{ij}^{1,1}$ captures the welfare gains from the entry of drug j alongside CGMs, $\Delta CS_{ij}^{1,0}$ represents the gains from product j had CGMs not been available, and $\Delta CS_{ij}^{0,1}$ represents the gains generated by CGMs had drug j not entered the market. In the last case, CGMs still affect insulin demand absent new drugs entering the market through ν_{ij}. The term Γ_{ij} measures the difference between the welfare gains from the two innovations and the unilateral contribution of each innovation, absent the other. A positive Γ_{ij} suggests that the welfare gains from product j and CGM entries together are greater than the sum of the gains from product j and those from CGMs unilateral decision, indicating synergies between the two innovations.

Figure 1.16 presents the average welfare gains across prescriptions from 2019 to 2021 and computes the average differential gains, $\overline{\Gamma}_j$ for the 36-hour and 42-hour products. The positive $\overline{\Gamma}_j$ for the 42-hour product suggests that the availability of CGMs amplifies the consumer welfare gains from this innovation. In contrast, $\overline{\Gamma}_j$ is negative for the 36-hour product, indicating that the technology does not similarly enhance its gains. The negative $\overline{\Gamma}_j$ for the 36-hour product is driven by the effect of the technology on competing pharmaceutical innovations and, in particular, the 42-hour product. Indeed, the average benefits from CGMs alone on the insulin market, $\overline{\Delta CS}_j^{0,1}$, are twice higher without the 36-hour drug than they are without the 42-hour product.

competitive pressure of having one more alternative in the market. New equilibrium prices for products still on the market can be computed from the Nash bargaining first-order conditions.

Figure 1.16: Daily consumer surplus, 2019-2021



Note: Each bar corresponds to the average consumer surplus in euros per day. For each product (horizontal axis), the first bar corresponds to the average consumer surplus from the joint entry of the product and CGMs, $\overline{\Delta CS}_j^{1,1}$; the second bar corresponds to the average surplus from the product entry, absent CGMs, $\overline{\Delta CS}_j^{1,0}$; the third bar is the average benefit from CGMs without this product, $\overline{\Delta CS}_j^{0,1}$; and the last bar is the average of the difference between the first bar and the sum of the last two, $\overline{\Gamma}_j$.

1.6.5 Towards product design

Innovative medical devices such as glucose sensors influence drug manufacturers' profits by creating new observable drug attributes that persistently impact perceived product differences. In France, the data generated by glucose sensors also enables drug manufacturers whose products excel in these new observable attributes to bargain higher prices with regulators, as the latter rewards pharmaceutical products for improving patient outcomes compared with existing alternatives through higher prices. As noted in the previous section, the information about drugs' performance provided by digital devices influences the benefits derived from the introduction of certain new drugs, which can impact incentives in drug development.

In this section, I document how the impact of digital medical devices on perceived drug differentiation can influence innovation incentives for insulin manufacturers. Specifically, I consider the patient-level information provided by CGMs during treatment choices and its potential effect on the characteristics of new drugs developed by manufacturers. When considering drug development decisions made by pharmaceutical companies, it's important to note that France is a small market unlikely to drive innovation on its own. This counterfactual assumes that the French market is representative of the global insulin market and that the influence of CGMs

is consistent across markets. I will discuss this assumption in more detail at the end of the section.

To explore the potential impact of CGMs on pharmaceutical innovation, I consider the perspective of an insulin manufacturer when deciding which new drug to develop from a range of potential products. I evaluate the expected profits from each possible product in environments both with and without CGMs, assuming that the manufacturer chooses to invest in the drug that offers the highest expected profits. This counterfactual analysis presumes that the drug manufacturer can anticipate future CGM adoption within the patient population during drug development. Specifically, I compute the expected profits from the entry of the 42-hour product, which effectively entered in 2018, and the expected profits from a fictive product that could be developed instead of the 42-hour one and could enter at the same time as the 42-hour drug did. Table 1.10 outlines the characteristics of the alternative product, called the '72-hour' drug. This product is designed to appear more or less appealing under incomplete and complete learning.[45] I assume that this product can be developed at a marginal cost (net of the cost shock, ζ_{jt}) 15% higher than that of the 42-hour product, that the regulatory agencies would approve it.[46] I limit the flows of profits to those earned during the first four years of market entry, from 2018 to 2021. I compute the market clearing prices with and without CGM and compare manufacturers' profits and consumer surplus in each scenario.

Figure 1.17 presents market share and firm profits under each scenario. The ranking of market shares and profits across products shifts depending on whether CGMs are available, suggesting that the information provided by CGMs can affect drug development decisions. Without CGMs, the 42-hour product captures a larger market share and is more profitable for the manufacturer than the 72-hour version is. However, when CGMs are introduced (and adopted by patients), the market share of the 72-hour version increases by 40%, and it becomes more profitable for the firm to launch the 72-hour product rather than the 42-hour one.

Figure 1.18, similar to Figure 1.16, presents the decomposition of consumer surplus using Equation (1.27). The consumer surplus from the 72-hour product combined with CGMs is 24% greater than that of the 42-hour product, driven by the 52% increase in consumer surplus owing to the joint entry of the drug and CGMs, as measured by Γ_j.

[45] The '72-hour' product features are designed to resemble the characteristics of the once-weekly insulin approved by the EMA in April 2024. Clinical trial outcomes are used to set μ_{nj}. However, the duration is 72 hours instead of 7 days to reduce the extent of out-of-sample extrapolation. In particular, the demand model does not capture the convenience of injecting insulin once a week instead of once a day. Hence, the 72-hour product is assumed to be injected once a day, similar to other available products. https://www.ema.europa.eu/en/medicines/human/EPAR/awiqli

[46] I also assume that the fixed costs associated with drug development are similar.

This last result suggests that the most profitable pharmaceutical innovation can shift on the basis of the technological environment in which demand occurs. By introducing new observable attributes for evaluating drug performance, complementary technologies, such as CGMs, can change the incentives for developing alternative products, potentially affecting future pharmaceutical innovations.

Discussion Innovation in pharmaceutical markets is driven by global demand. As mentioned earlier, France is not large enough to drive innovation. The insights from this counterfactual exercise rely on the external validity of my estimates for other geographical markets, which I discuss briefly here. The econometric model estimates preferences for insulin products and the primitives of the price setting model. On the insulin demand side, given the adoption of glucose sensors, the results suggest the device generates new observable patient-specific attributes and limited information spillovers to nonusers. This outcome is likely to hold in contexts beyond France. However, the contribution of ν_{ij} is heterogeneous across patients. The aggregate effect of observing ν_{ij} depends on its distribution in the whole patient population and among patients adopting the device. Hence, the magnitude of the effect may vary.

Considering the price setting model, the regulator in France acts as a monopsony in a small country. This feature affects the external validity of my results for contexts like the US. In more fragmented settings where drug manufacturers bargain with insurers, the disagreement payoffs will affect equilibrium prices and the response of drug pricing to CGM information. Yet, the consumer surplus generated by a given prescription drug also matters for private insurers. They may also be inclined to pay higher prices for drugs that result in a higher match value.

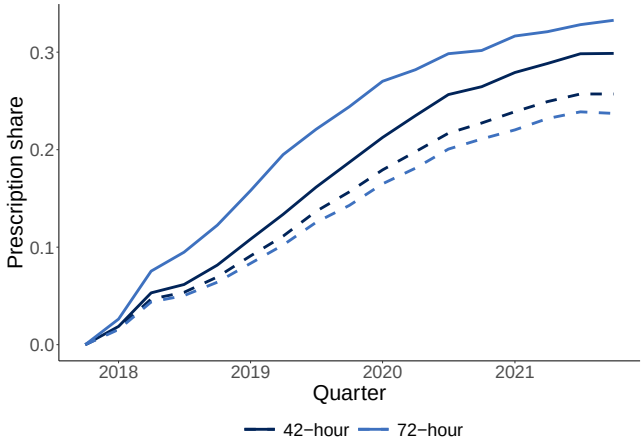
Last, in my setting, CGM adoption is taken as given. The results suggest little social learning and the effect comes from insulin demand for glucose sensor users. Hence, the extent and composition of CGM adoption matter. Health insurers' decisions to cover CGMs and under which conditions are beyond the scope of this paper.

Table 1.10: Product features

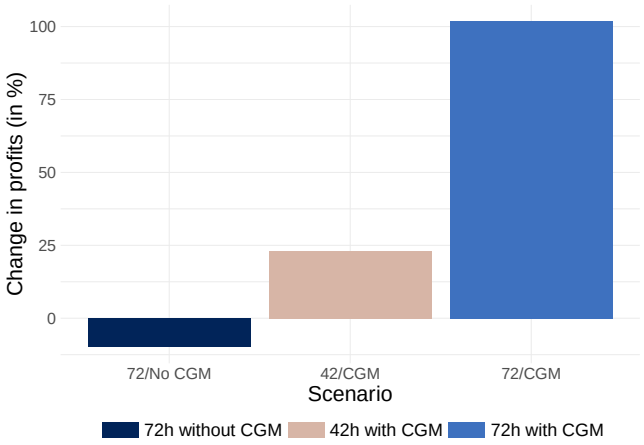
	Supply		Priors		μ_{nj}						
	mc_j	Duration	μ_{0j}	$(V_j^0)^{1/2}$	μ_{1j}	μ_{2j}	μ_{3j}	μ_{4j}	μ_{5j}	μ_{6j}	μ_{7j}
42-hour (Obs)	mc_{42}	1.75	-3.49	2.65	0.94	0.5	0.40	0.26	0.31	0.77	0.47
72-hour	$1.15 \times mc_{42}$	3	-3.57	2.70	$0.5\mu_{n,42}$			0.25		$1.1\mu_{n,42}$	

Note: The 72-hour product is inspired by the clinical trial outcomes for insulin icodec, approved by the EMA in April 2024. The marginal cost is assumed to be 15% higher than the 42h version.

Figure 1.17: 42-hour and 72-hour product entry



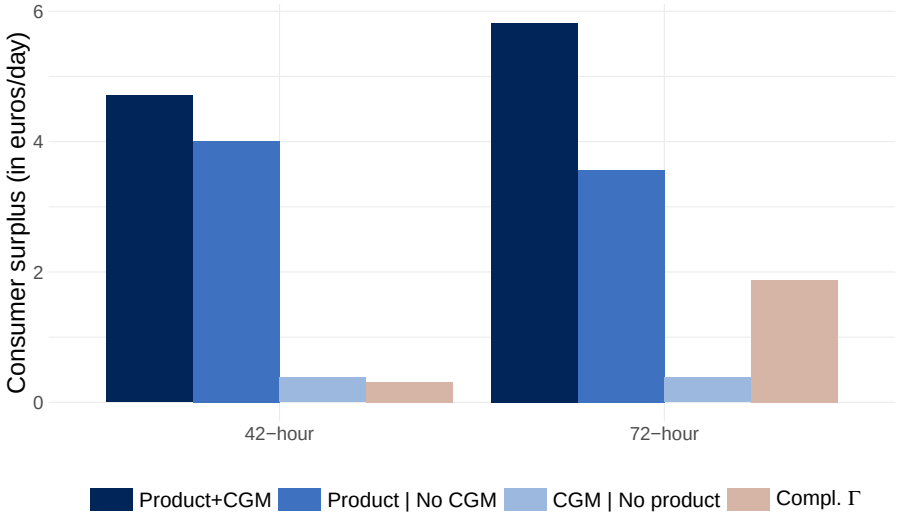
(a) Market shares, with and without (dashed) CGMs



(b) Firm's profit from product entry

Note: Figure 1.17a compares the average choice probabilities for each product over time following entry. The dashed lines correspond to product shares in a market that does not provide access to CGMs, whereas the plain lines consider the availability of CGMs for patients. Figure 1.17b plots the changes in profits compared with the profits from the entry of the 42-hour product in an environment without CGMs. The first bar corresponds to the difference in profits if the 72-hour product enters an environment without CGMs. The second bar corresponds to the profit difference if the 42-hour product enters an environment with CGMs, and the third bar corresponds if the 72-hour product enters an environment with CGMs.

Figure 1.18: Daily consumer surplus, 2019-2021



Note: Each bar corresponds to a measure of average consumer surplus in euros per day. For each product (horizontal axis), the first bar corresponds to the average consumer surplus from the joint entry of the product and CGMs, $\overline{\Delta CS}_j^{1,1}$; the second bar corresponds to the average surplus from the product entry, absent CGMs, $\overline{\Delta CS}_j^{1,0}$; the third bar is the average benefit from CGMs without this product, $\overline{\Delta CS}_j^{0,1}$; and the last bar is the average of the difference between the first bar and the sum of the last two, $\overline{\Gamma}_j$.

1.7 Conclusion

This paper studies how digital medical devices affect pharmaceutical demand in the context of CGM insights for insulin choice. To that end, I develop a tractable model of demand and supply for insulin embedding: (i) patient-specific learning about treatment performance through the CGM, (ii) dynamic physician-level learning about new drugs from patient experiences, and (iii) price setting by pharmaceutical companies and the regulator, both of which internalize demand-side learning. The model is estimated using medical claims data from France that records insulin prescriptions by diabetes specialists.

The results from the structural estimation highlight how insights from digital health data reshaped physicians' preferences when prescribing insulin to diabetic patients. Despite significant information frictions about new products, affecting both CGM users and nonusers initially, consumer welfare gains mostly accrued to patients using the device, while the benefits to nonusers were ten times smaller. The technology had a heterogeneous effect across new products. Physicians learn faster about new drugs that present features relevant to overcoming the flaws of existing drugs revealed by the technology. By inducing CGM users to switch treatment away from old molecules to these new drugs, the device generated opportunities to learn about the quality of the drug that would not have occurred without the technology. As CGMs did not improve the quality of the experience signals received by the physicians about new drugs and the number of learning opportunities remained fixed, physicians would have learned more about some products without the device. Assuming that CGMs allows for complete learning about patient-insulin clinical match values, this framework enables the measurement of the losses, in terms of foregone profits and consumer surplus, from relying on partial information when deciding on treatment. This paper also highlights the cross-market complementarities between medical devices and pharmaceuticals, exploring the potential long-term effects of CGMs on the insulin market. Specifically, the most profitable pharmaceutical innovation strategy can vary depending on the technological environment in which insulin choices are made, as CGMs introduce new observable attributes for evaluating drug performance that enter pharmaceutical demand.

This study opens up several important avenues for future research. First, estimating the patient-product clinical match values does not separately identify the preference for the effects on the average and glucose profile from their level, as they remain unobserved in claims data. Further separating these criteria would shed light on whether the old attributes are phased out

by the new attributes generated from CGMs. Second, the analysis focuses on the impact of CGMs on existing insulin patients. It does not account for the potential benefits for new insulin users, such as improved diagnosis matching by physicians at therapy initiation (Crawford and Shum (2005)). Third, the framework is estimated for France, where the regulator bargains for the price of prescription drugs with insulin manufacturers and provides complete coverage for insulin and CGM expenses through the universal health insurance scheme. This leads to low prices and inelastic demand. The impact of CGMs may vary in environments where prices adjust more freely, and cost differences influence insulin choices. Finally, this project focuses on cross-market externalities, and a complete welfare analysis of CGMs is beyond its scope. Since CGMs increase monitoring costs relative to traditional devices, broad coverage significantly increases healthcare expenses. Identifying which patients benefit most from CGM data and how they benefit would be crucial for designing an optimal coverage strategy (Chandra and Skinner (2012), Conner et al. (2024)). Their impact on the organization of care also presents important avenues for research that could inform policy debates, as healthcare systems such as the NHS in the UK consider adopting these technologies to support system reform.[47]

Bibliography

Agarwal, N., Moehring, A., Rajpurkar, P., and Salz, T. (2023). Combining human expertise with artificial intelligence: Experimental evidence from radiology. Technical report, National Bureau of Economic Research.

Agha, L. (2014). The effects of health information technology on the costs and quality of medical care. *Journal of health economics*, 34:19–30.

Alsan, M., Durvasula, M., Gupta, H., Schwartzstein, J., and Williams, H. (2024). Representation and extrapolation: evidence from clinical trials. *The quarterly journal of economics*, 139(1):575–635.

Barahona, N., Otero, C., and Otero, S. (2023). Equilibrium effects of food labeling policies. *Econometrica*, 91(3):839–868.

Bergemann, D. and Välimäki, J. (2006). Dynamic pricing of new experience goods. *Journal of Political Economy*, 114(4):713–743.

[47] https://www.independent.co.uk/news/uk/politics/nhs-smartwatches-diabetes-streeting-labour-b2632165.html accessed on October 25th, 2024.

Bresnahan, T. F. and Trajtenberg, M. (1995). General purpose technologies 'engines of growth'? *Journal of econometrics*, 65(1):83–108.

Brown, Z. Y. and Jeon, J. (2024). Endogenous information and simplifying insurance choice. *Econometrica*, 92(3):881–911.

Chandra, A. and Skinner, J. (2012). Technology growth and expenditure growth in health care. *Journal of Economic Literature*, 50(3):645–680.

Chehregosha, H., Khamseh, M. E., Malek, M., Hosseinpanah, F., and Ismail-Beigi, F. (2019). A view beyond hba1c: role of continuous glucose monitoring. *Diabetes Therapy*, 10:853–863.

Ching, A. T. (2010a). Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration. *International Journal of Industrial Organization*, 28(6):619–638.

Ching, A. T. (2010b). A dynamic oligopoly structural model for the prescription drug market after patent expiration. *International Economic Review*, 51(4):1175–1207.

Conner, P., Einav, L., Finkelstein, A., and Persson, P. (2024). Targeting precision medicine: Evidence from prenatal screening. *Journal of Political Economy*.

Coscelli, A. and Shum, M. (2004). An empirical model of learning and patient spillovers in new drug entry. *Journal of econometrics*, 122(2):213–246.

Crawford, G. S. and Shum, M. (2005). Uncertainty and learning in pharmaceutical demand. *Econometrica*, 73(4):1137–1173.

Currie, J. M. and MacLeod, W. B. (2020). Understanding doctor decision making: The case of depression treatment. *Econometrica*, 88(3):847–878.

Dahlstrand, A. (2024). Defying distance? the provision of services in the digital age.

Dickstein, M. J. (2021). *Efficient provision of experience goods: Evidence from antidepressant choice*.

Doraszelski, U., Lewis, G., and Pakes, A. (2018). Just starting out: Learning and equilibrium in a new market. *American Economic Review*, 108(3):565–615.

Dranove, D., Forman, C., Goldfarb, A., and Greenstein, S. (2014). The trillion dollar conundrum: Complementarities and health information technology. *American Economic Journal: Economic Policy*, 6(4):239–270.

Dranove, D., Garthwaite, C., and Hermosilla, M. (2022). Does consumer demand pull scientifically novel drug innovation? *The RAND Journal of Economics*, 53(3):590–638.

Dubé, J.-P., Hitsch, G. J., and Rossi, P. E. (2010). State dependence and alternative explanations for consumer inertia. *The RAND Journal of Economics*, 41(3):417–445.

Dubois, P., Gandhi, A., and Vasserman, S. (2022). Bargaining and international reference pricing in the pharmaceutical industry. Technical report, National Bureau of Economic Research.

Dubois, P., Griffith, R., and O'Connell, M. (2018). The effects of banning advertising in junk food markets. *The Review of Economic Studies*, 85(1):396–436.

Dubois, P. and Tunçel, T. (2021). Identifying the effects of scientific information and recommendations on physicians' prescribing behavior. *Journal of health economics*, 78:102461.

Einav, L., Finkelstein, A., and Polyakova, M. (2018). Private provision of social insurance: drug-specific price elasticities and cost sharing in medicare part d. *American Economic Journal: Economic Policy*, 10(3):122–153.

Eizenberg, A. (2014). Upstream innovation and product variety in the us home pc market. *Review of Economic Studies*, 81(3):1003–1045.

Erdem, T. and Keane, M. P. (1996). Decision-making under uncertainty: Capturing dynamic brand choice processes in turbulent consumer goods markets. *Marketing science*, 15(1):1–20.

Gotham, D., Barber, M. J., and Hill, A. (2018). Production costs and potential prices for biosimilars of human insulin and insulin analogues. *BMJ global health*, 3(5):e000850.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review*, 103(1):145–177.

Grennan, M., Myers, K., Swanson, A., and Chatterji, A. (2024). No free lunch? welfare analysis of firms selling through expert intermediaries. Technical report.

Grennan, M. and Town, R. J. (2020). Regulating innovation with uncertain quality: information, risk, and access in medical devices. *American Economic Review*, 110(1):120–61.

Guerci, B., Roussel, R., Levrat-Guillen, F., Detournay, B., Vicaut, E., De Pouvourville, G., Emery, C., and Riveline, J.-P. (2023). Important decrease in hospitalizations for acute diabetes events following freestyle libre system initiation in people with type 2 diabetes on basal insulin therapy in france. *Diabetes Technology & Therapeutics*, 25(1):20–30.

Handel, B. and Kolstad, J. (2017). Wearable technologies and health behaviors: New data and new methods to understand population health. *American Economic Review*, 107(5):481–485.

Handel, B. R. and Kolstad, J. T. (2015). Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8):2449–2500.

Handel, B. R. and Misra, K. (2015). Robust new product pricing. *Marketing Science*, 34(6):864–881.

Hitsch, G. J. (2006). An empirical model of optimal dynamic product launch and exit under demand uncertainty. *Marketing Science*, 25(1):25–50.

Igami, M., Kusaka, S., Qiu, J., and Tran, T. L. (2024). Welfare gains from product & process innovations: The case of lcd panels, 2001–2011.

Jin, G. Z. and Leslie, P. (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards. *The Quarterly Journal of Economics*, 118(2):409–451.

Jin, Y. and Vasserman, S. (2021). Buying data from consumers: The impact of monitoring programs in us auto insurance. Technical report, National Bureau of Economic Research.

Lee, J., McCullough, J. S., and Town, R. J. (2013). The impact of health information technology on hospital productivity. *The RAND Journal of Economics*, 44(3):545–568.

Litvinova, O., Eitenberger, M., Bilir, A., Yeung, A. W. K., Parvanov, E. D., MohanaSundaram, A., Horbańczuk, J. O., Atanasov, A. G., and Willschke, H. (2023). Patent analysis of digital sensors for continuous glucose monitoring. *Frontiers in Public Health*, 11:1205903.

Maini, L., Feng, J., Hwang, T., and Klimek, J. (2022). Biosimilar entry and the pricing of biologic drugs. *Available at SSRN 3760213*.

McCullough, J. S., Parente, S. T., and Town, R. (2016). Health information technology and patient outcomes: the role of information and labor coordination. *The RAND Journal of Economics*, 47(1):207–236.

Patel, M. S., Asch, D. A., and Volpp, K. G. (2015). Wearable devices as facilitators, not drivers, of health behavior change. *Jama*, 313(5):459–460.

Petrin, A. (2002). Quantifying the benefits of new products: The case of the minivan. *Journal of political Economy*, 110(4):705–729.

Roberts, J. H. and Urban, G. L. (1988). Modeling multiattribute utility, risk, and belief dynamics for new consumer durable brand choice. *Management Science*, 34(2):167–185.

Seaquist, E., Ritzel, R., Pettus, J., and Khunti, K. (2017). Session 1: Basal insulins: Still innovating after all these years. *DIABETES*.

Shapiro, C. (1983). Optimal pricing of experience goods. *The Bell Journal of Economics*, pages 497–507.

Shields, A. and Sankaranarayanan, S. (2016). Basal insulin regime change from lantus to toujeo resulted in fewer hypoglycaemic episodes in a 28-year-old man with diabetes mellitus. *Case Reports*, 2016:bcr2016215831.

Siamashvili, M., Davis, H. A., and Davis, S. N. (2021). Nocturnal hypoglycemia in type 1 and type 2 diabetes: an update on prevalence, prevention, pathophysiology and patient awareness. *Expert Review of Endocrinology & Metabolism*, 16(6):281–293.

Trajtenberg, M. (1989). The welfare analysis of product innovations, with an application to computed tomography scanners. *Journal of political Economy*, 97(2):444–479.

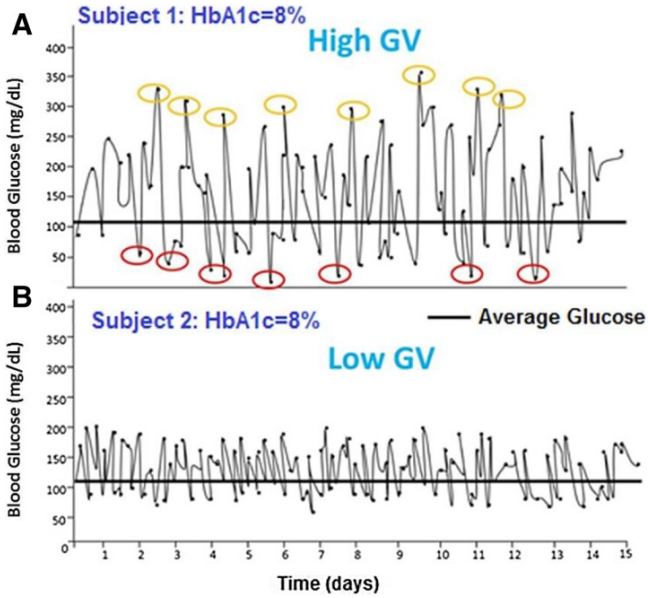
Tunçel, T. (2024). Should we prevent off-label drug prescriptions? empirical evidence from france. *Review of Economic Studies*, page rdae060.

Zeltzer, D., Einav, L., Rashba, J., and Balicer, R. D. (2024). The impact of increased access to telemedicine. *Journal of the European Economic Association*, 22(2):712–750.

Zhu, Z. (2023). The value of patients: Heterogenous physician learning and generic drug diffusion. Technical report, HEDG, c/o Department of Economics, University of York.

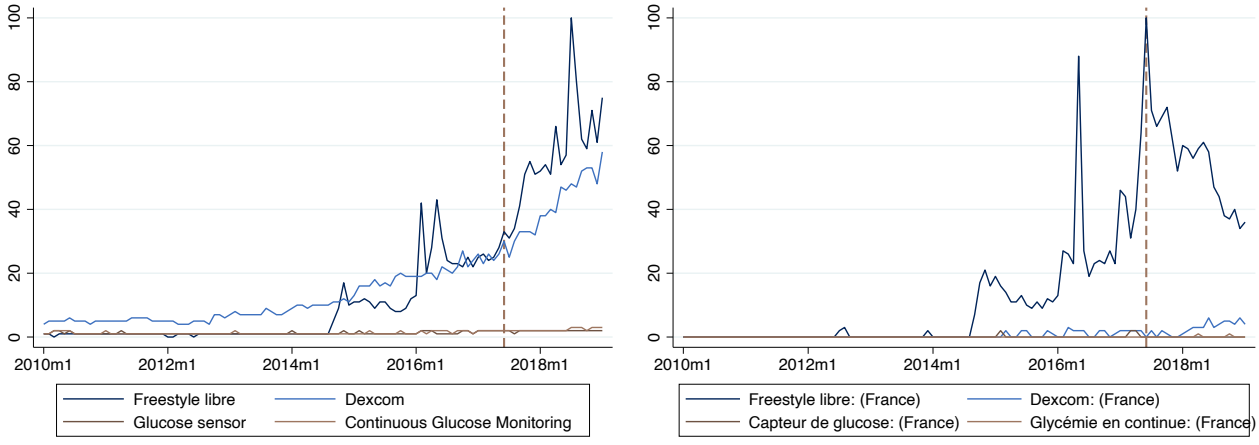
1.8 Additional figures and tables

Figure 1.19: Glucose variability



Note: Different patterns of glycemic variability (GV) in two patients with same hemoglobin A1C (HbA1c). 15-day glucose traces of two patients who had identical HbA1c of 8.0% but different degrees of GV. High GV in patient 1 was reflected by numerous episodes of both hypo- and hyperglycemia (a), whereas low GV in subject 2 resulted in no such episodes (b). Patient 1 (a) had visibly higher glucose fluctuations than patient 2 (b) that resulted in seven episodes of moderate hypoglycemia ($\leq$ 50mg/dL) and eight episodes of moderate hyperglycemia ($\geq$ 350mg/dL). Reproduced from Chehregosha et al. (2019) under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/). No changes were made to the original figure.

Figure 1.20: Google trends for Continuous Glucose Monitoring, 2010-2019

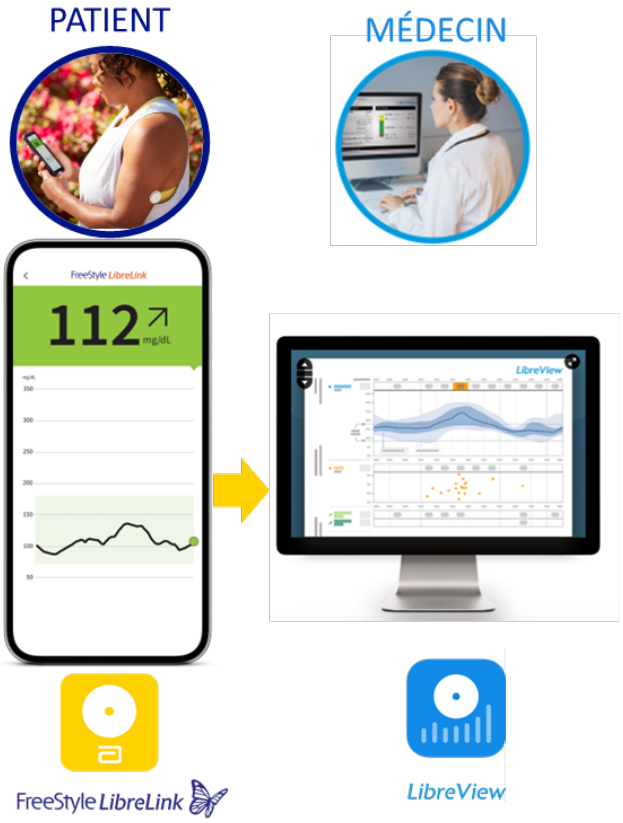


(a) All countries

(b) France

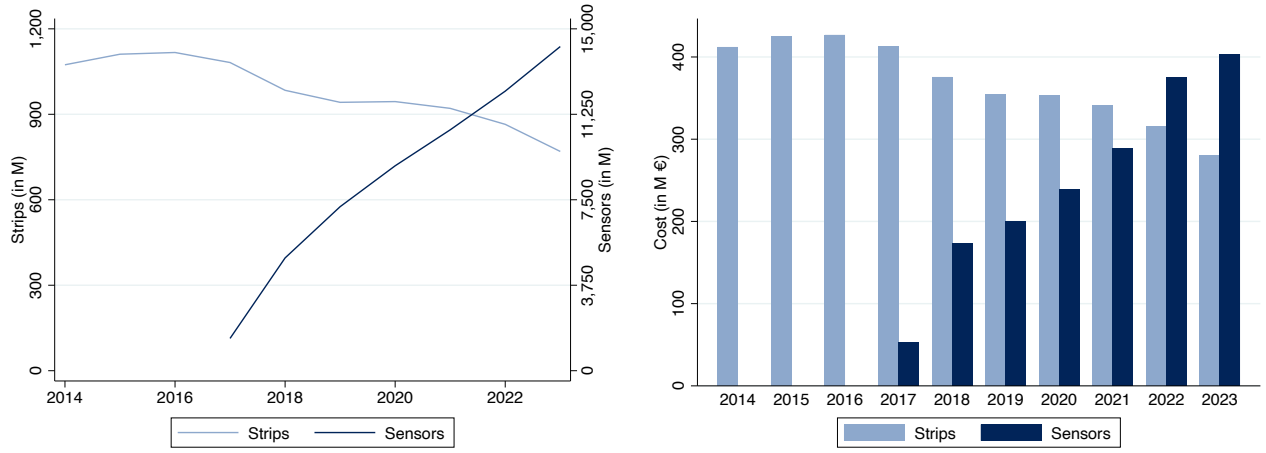
Note: In each figure, the legend corresponds to the keywords. In Figure 1.20b, 'Capteur de glucose' means 'glucose sensor' and 'Glycémie en continue' corresponds to 'Continuous glucose monitoring'. Freestyle Libre is the technology developed by Abbott in 2014. The former technology developed by Abbott - Freestyle Navigator II (not displayed) - faced a low index for the whole time it was available. Dexcom is also developing CGM.

Figure 1.21: Continuous Glucose Monitoring



Note: L'application FreeStyle LibreLink permet, si vous le souhaitez, le partage automatique et à distance de vos données de glucose avec votre médecin à travers la plateforme d'analyse en ligne sécurisée LibreView. Ce partage automatique nécessite que vous ayez au préalable créé un compte LibreView, puis que vous vous connectiez au cabinet de votre médecin.Avec un compte LibreView, vous pouvez : (1) Partager automatiquement vos données de glucose avec votre médecin. (2) Votre médecin pourra ainsi accéder à vos données de glucose depuis son compte LibreView, sans que vous ayez besoin de vous rendre à son cabinet. (3) Retrouver des rapports synthétiques et intuitifs qui mettent en évidence votre profil et vos tendances de variation de vos taux de glucose. (4) Accéder à tout moment à toutes vos données, stockées en toute sécurité sur LibreView. Illustration from Abbott's website, used in accordance with Abbott's copyright policy for non-commercial purposes. Source: Abbott website accessed on October 10th, 2024.

Figure 1.22: Glucose monitoring volumes and value in France, 2014-2023

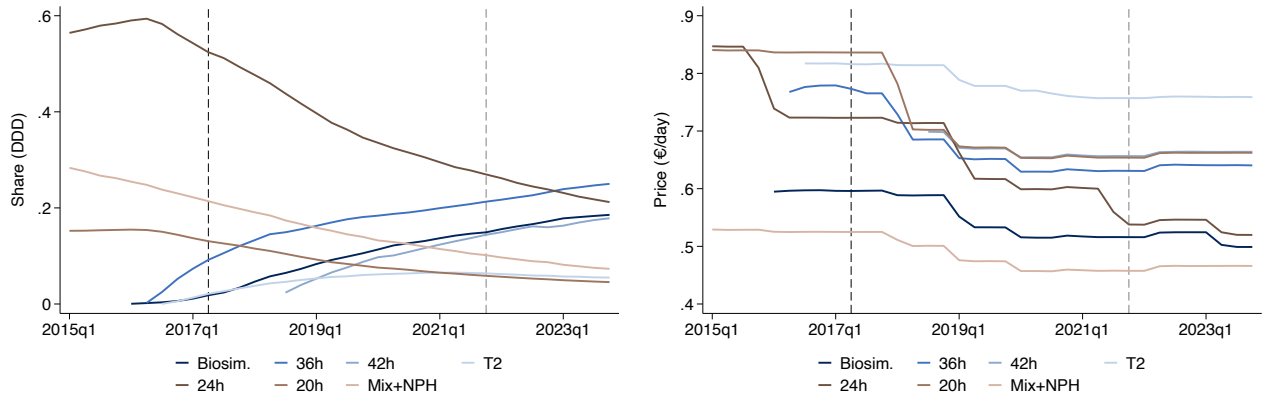


(a) Number of tests, in million

(b) Value, in million euros

Note: Test volume and value for all patients in France. In Figure 1.22a, the number of tests performed via sensors is computed from the number of sensors and frequency of automatic measurement. In Figure 1.22b, strip costs do not include the lancing device cost. Source: Open LPP.

Figure 1.23: Aggregate insulin sales and prices, 2015-2023

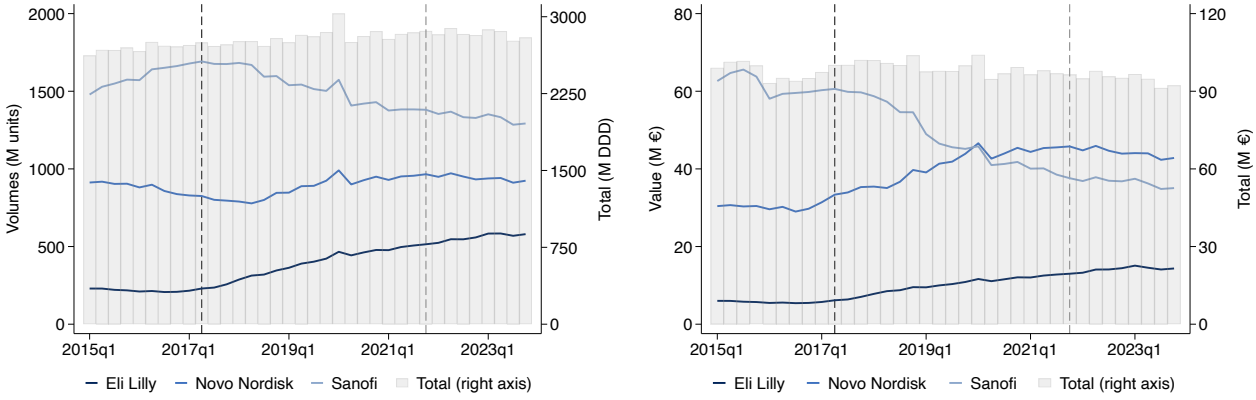


(a) Units reimbursed

(b) Price per daily consumption

Note: The first vertical line represents the start of CGM coverage. The second line corresponds to the end of the period considered in the analysis. Figure 1.23a plots the share for each product over time as a percentage of all insulin units reimbursed. Figure 1.23b plots the daily price, assuming 20 units per day. The 36h daily dose price is adjusted to account for the increased units required when switching from the regular 24h.

Figure 1.24: Sales volume and value, by manufacturer, 2015-2023

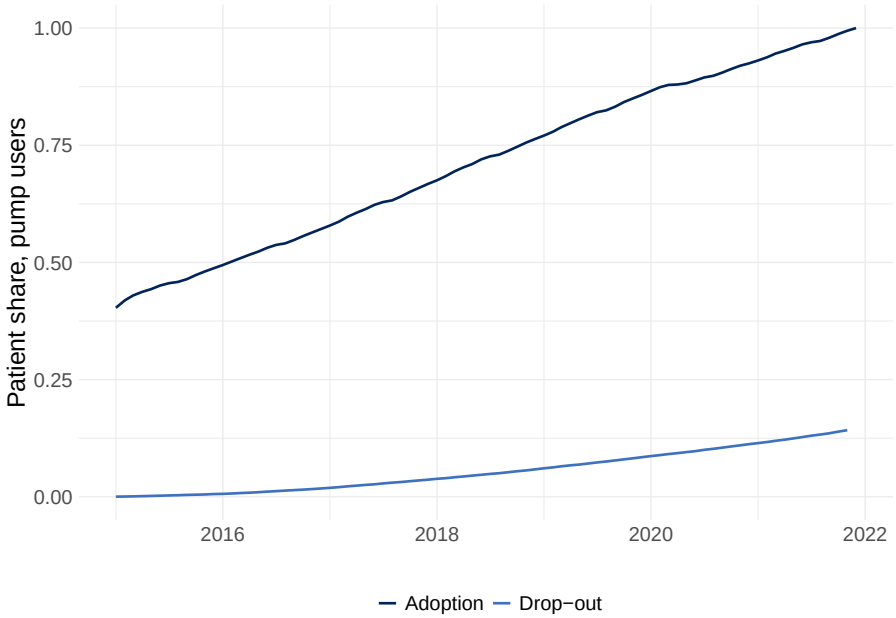


(a) Units reimbursed, million units

(b) Sales value, million €

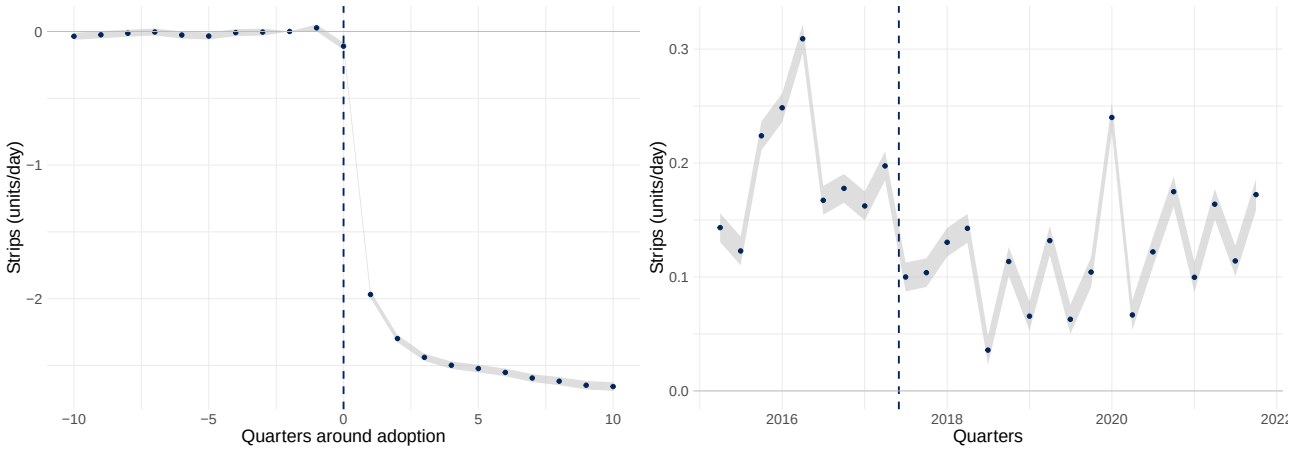
Note: The first vertical line represents the start of CGM coverage. The second line corresponds to the end of the period considered in the analysis.

Figure 1.25: Insulin pump usage, CDF



Note: This figure plots the cumulative distribution function of the first and last pump reimbursement date for patients relying on an insulin pump between 2015 and 2021. These distributions present no discontinuity/change in trends after the CGM coverage decision.

Figure 1.26: Strips reimbursement per patient

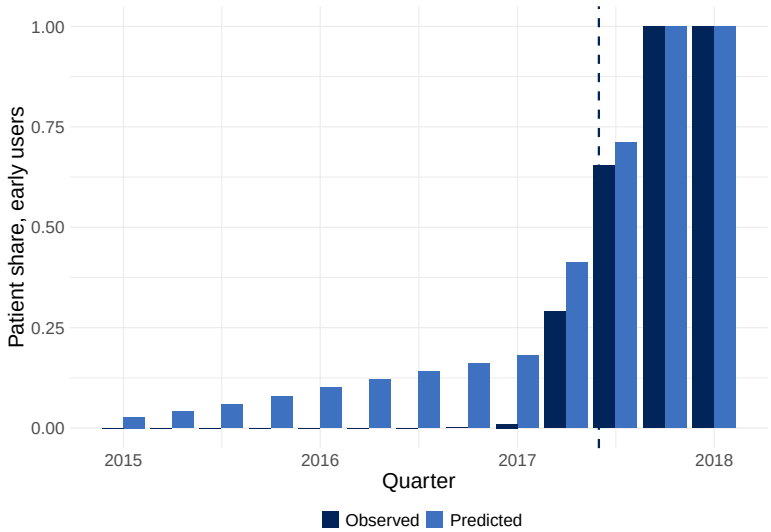


(a) Before-after CGM adoption

(b) Non-users over time

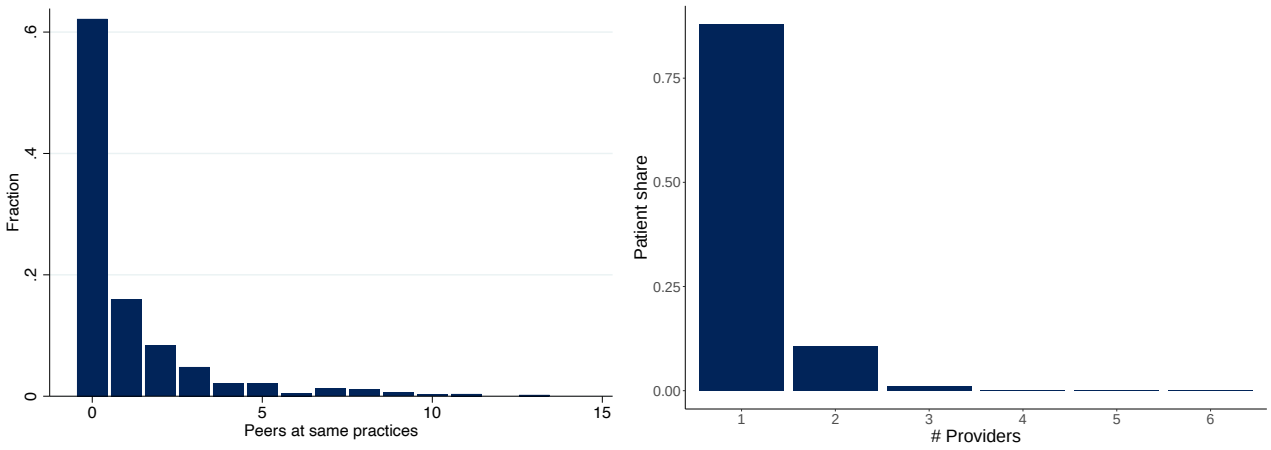
Note: Figure 1.26a presents the estimates from an event study model considering the number of strips reimbursed in the quarters around CGM adoption. CGM adoption is identified from a patient's first prescription of a CGM. The regression includes patient fixed-effects and focuses on patients with their first CGM prescription from January 2018 onwards. Figure 1.26b plots the linear regression coefficients of strip reimbursement over time for nonusers.

Figure 1.27: Observed and predicted CGM adoption date for 'early' users



Note: Figure 1.27 compares the first CGM prescription date observed in the data to the one predicted by the model summarized in Table 1.13 for 'early adopters'. 'Early adopters' are patients whose first CGM reimbursement happens before January 2018 such that the first CGM prescription may not coincide with their adoption date. 40k patients are adopting before January 2018, among which 7k face a predicted adoption date different from their observed date.

Figure 1.28: Diabetes specialists outside of hospital practices

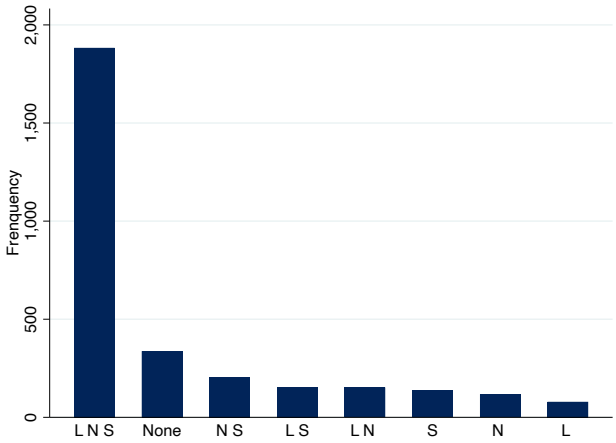


(a) Number of "peers"

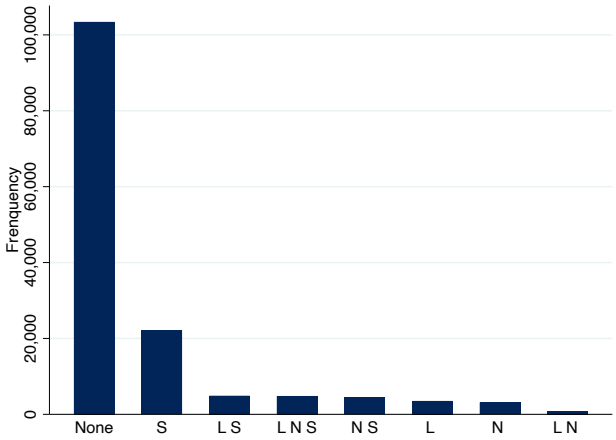
(b) Number of practitioners by patient

Note: Figure 1.28a represents the number of "peers" working at the same practice for diabetes specialists working outside the hospital. 62% are working in an environment without any other diabetes specialist, 78% with at most one peer. Figure 1.28a uses data from the practionners directory in France available at https://annuaire.sante.fr/web/site-pro/extractions-publiques. Figure 1.28b plots the number of diabetes specialists working outside the hospital by patient.

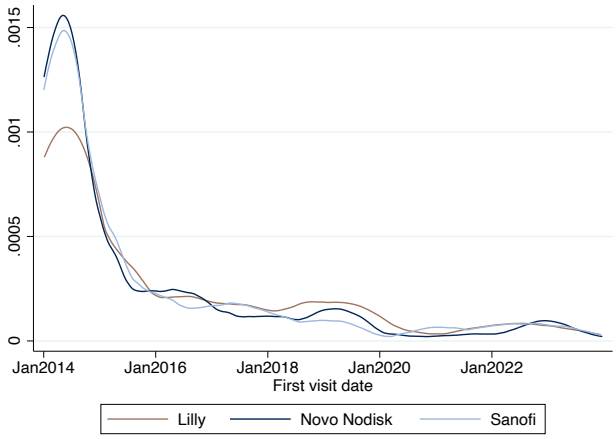
Figure 1.29: Insulin manufacturer detailing to physicians, 2014-2023



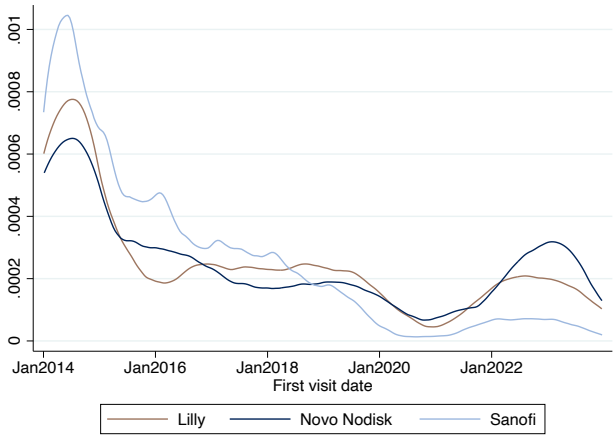
(a) Specialists and manufacturers interactions



(b) GPs and manufacturers interactions



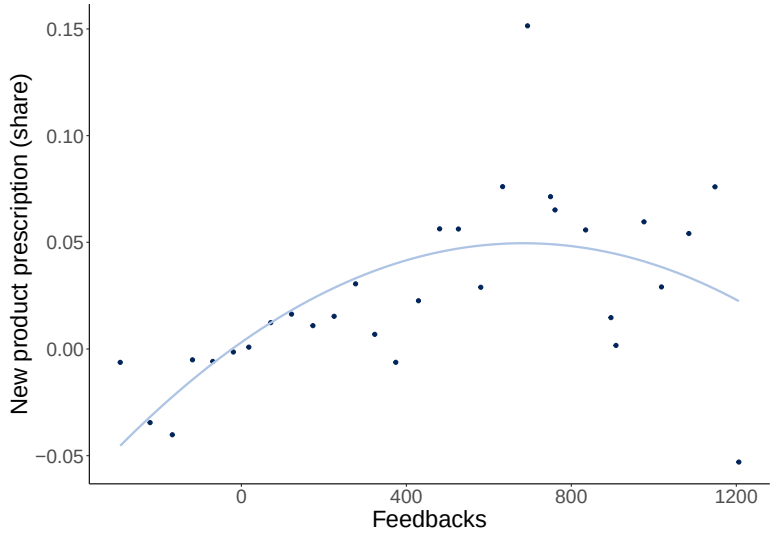
(c) First detailing to specialists



(d) First detailing to GPs

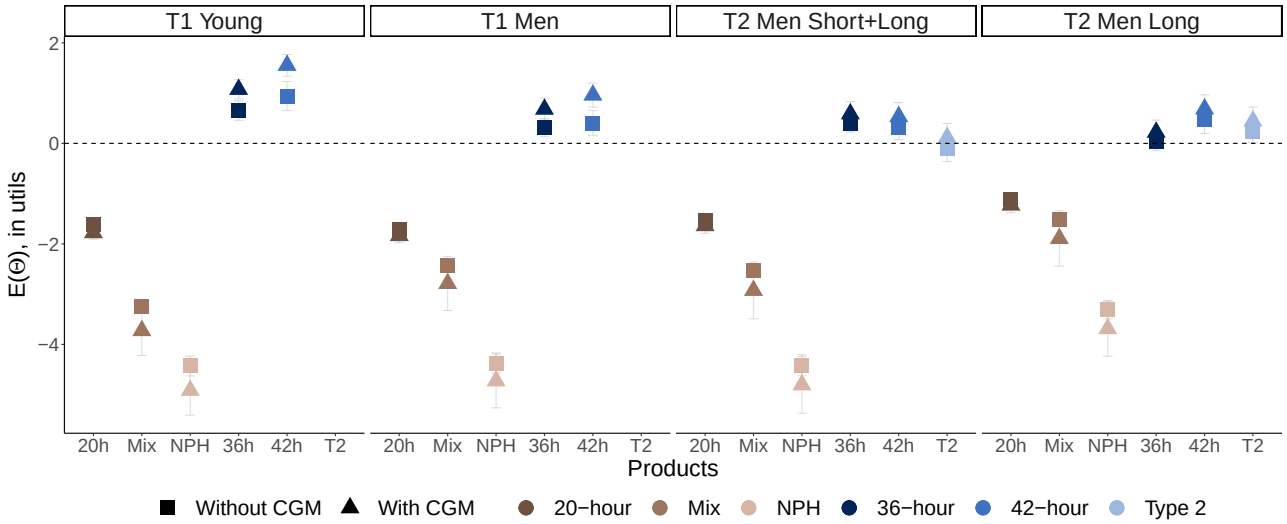
Note: Descriptive evidence of pharmaceutical detailing by insulin manufacturers from 2014 to 2023 are provided in Figure 1.29. Diabetes specialists are more likely to interact with insulin manufacturers than GPs. Over the period, 90% of specialists interacted at least once with a pharmaceutical company, and 60% of them can be linked to the three manufacturers. They interact on average once per year and, in most cases, since the beginning of my sample period. On the other hand, 70% of GPs never interacted with either one of these companies. While 25% interacted with Sanofi, the manufacturer with the most diverse prescription drug portfolio, less than 10% interacted with Eli Lilly and/or Novo Nordisk, which have a less diverse portfolio.

Figure 1.30: Physician prescription share and the information set size, city specialist



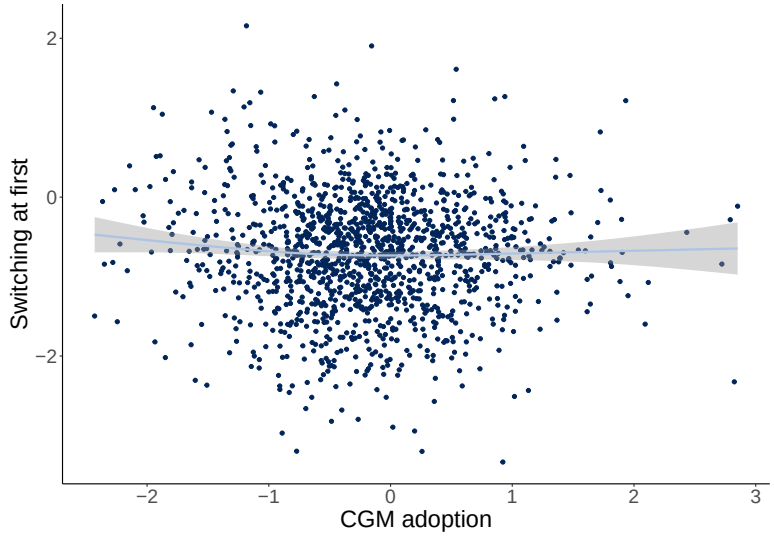
Note: The horizontal axis corresponds to the amount of feedback from real-life experience with the new product j received by the physician up to the previous period. The vertical axis corresponds to product j prescription share for patients not using the product before the appointment. For each variable, I consider the residuals from a linear model controlling for physician fixed effect, product-specific quarter fixed effect, and the average demographics of the patient visiting physician k in a given period. The figure focuses on prescriptions by diabetes specialists working outside the hospital and excludes the new Type 2 product directed towards a subset of patients.

Figure 1.32: Perceived match value with/without a CGM



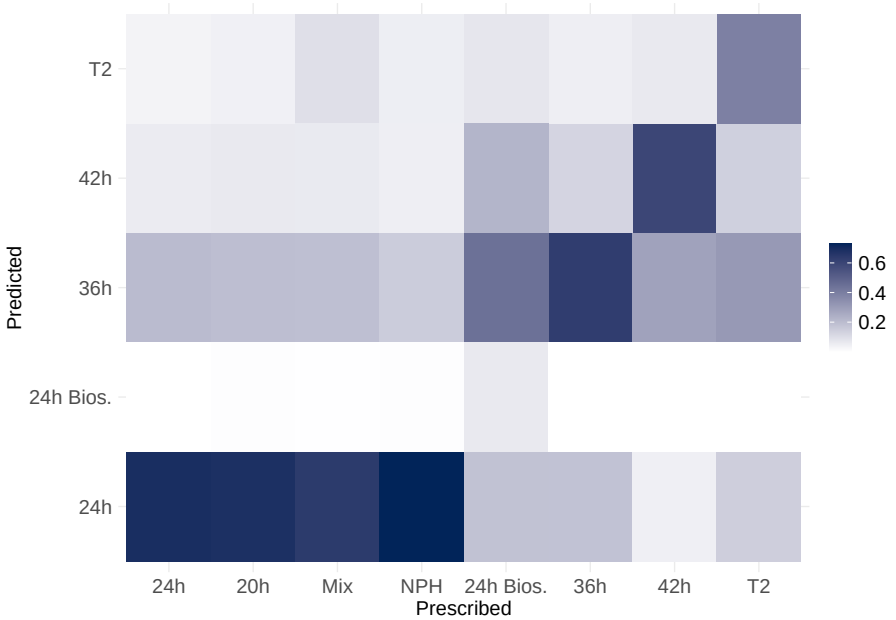
Note: This figure presents the perceived match value without the technology (squares) and with CGM (triangles) in utils (vertical axis) across products (horizontal axis). The perceived match value without the technology corresponds to μ_{nj} while the perceived match value with the technology is $\mu_{nj} + \nu_{nj}$ where ν_{nj} is the average across patients within a cluster.

Figure 1.31: Physician-level heterogeneity



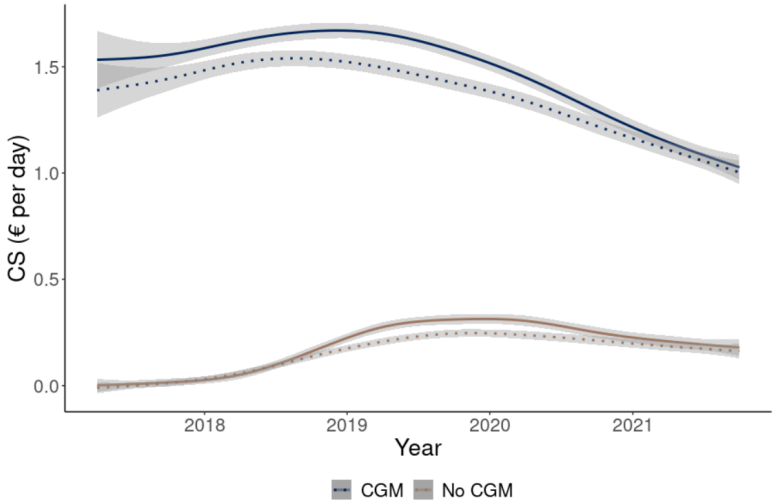
Note: This figure presents the physician-level propensity to see patients with CGMs (horizontal axis) against her propensity to switch patients in the first appointment after they adopted CGMs (vertical axis). Each propensity corresponds to the parameter of a logistic regression where the outcome variable is 'CGM adoption' (horizontal axis) and 'Switching insulin in the first appointment with CGM' (vertical axis), and patients' characteristics are included as controls. One point corresponds to one physician. The fitted line corresponds to a flexible non-parametric smoothing of the data with a 95% confidence interval.

Figure 1.33: Prescribed vs predicted treatment choice



Note: This figure compares the product prescribed in the actual data j^{obs} to $\hat{j} = \arg\max \hat{\bar{u}}_{ikjv}$. The idiosyncratic shock is excluded from the prediction. Note that $\hat{j}$ never corresponds to the 20-hour, insulin mixes nor human insulin treatments.

Figure 1.34: Consumer surplus under observed allocation and after sensor reallocation



Note: The figure presents the compensating variation (vertical axis), in euros per day, following a prescription happening in period t (horizontal axis). The average consumer welfare is presented for patients with CGM (blue curve) and without (brown curve). The plain line represents consumer welfare from the observed adoption pattern. Dotted lines present welfare after reallocating sensors across patients.

Table 1.11: Summary statistics, patient level

	(1)	(2)	(3)	(4)
	All	Type I	Type II	
			Long+short	Long only
N ('000)	333.4	94.2	133.4	105.8
Age	57	49	60	62
Female	0.434	0.405	0.449	0.440
Type 1	0.283	1.000	0.000	0.000
Low-income	0.126	0.135	0.126	0.117
Residential area				
Deprivation index	0.262	0.044	0.336	0.371
Population('000)	41.0	44.1	41.2	37.7
Chronic conditions				
Hypertension	0.692	0.442	0.783	0.799
Hypercholesterolemia	0.619	0.401	0.700	0.712
Analgesics	0.412	0.286	0.474	0.446
Obesity	0.324	0.169	0.412	0.351
Cardiovascular	0.318	0.205	0.382	0.338
Anxiolytics	0.174	0.143	0.193	0.177
Antidepressants	0.168	0.135	0.192	0.167
Respiratory	0.148	0.103	0.178	0.150
Hypnotics	0.108	0.078	0.127	0.110
Cancer	0.097	0.064	0.115	0.104
Neuroleptics	0.043	0.036	0.045	0.046
Dialyse	0.016	0.013	0.022	0.010
Prescriptions	7	8	8	5
CGM Users	0.449	0.728	0.531	0.096
CGM Temporary	0.073	0.050	0.080	0.174
Pump Users	0.067	0.171	0.041	0.008
Insulin switch	0.548	0.579	0.616	0.435
Nb switches	1.446	1.437	1.538	1.294

Note: The sample is restricted to patients between 18 and 75 in 2015 who had already used long-acting insulin in early 2016, who had gotten a prescription from a diabetes specialist, and who did not rely exclusively on an insulin pump over the sample period. The Deprivation index is computed based on 2015 measures of unemployment, blue-collar workers, high school graduates shares and the median income by consumption unit at the city level from the national statistic institute (INSEE). It is centered around zero, goes from -6.1 to 10.3, and the variance is 2.72. Negative values stand for more favorable areas. In 2015, the median individual in France lived in a 9,423 inhabitants city, and the deprivation index is around 0.116. The number of prescriptions is restricted to prescriptions written by diabetes specialists. The number of insulin switches is computed on the sample of patients who switched at least once. Appendix 1.9.1 details the sample construction.

Table 1.12: Summary statistics, physician level

	(1) All	(2) Specialists	(3) Hospital	(4) GP	(5) Care Center	(6) Entering non-GP
N	108,542	845	1,022	99,244	7,062	369
N (Share)		0.008	0.009	0.914	0.065	0.003
Patients (#)	5	122	179	5	3	10
Prescriptions (#)	22	739	428	23	5	17
Prescriptions (Share)		0.089	0.164	0.715	0.028	0.004
Ever CGM (%)	0.581	0.946	0.994	0.577	0.526	0.691
CGM users	2	62	77	2	1	6
CGM prescription (#)	1	152	72.5	1	1	5
CGM (Share)		0.134	0.231	0.592	0.031	0.012
	A. Ever prescribed (%)					
All products	0.084	0.670	0.886	0.070	0.082	0.263
20-hour	0.485	0.983	0.994	0.482	0.389	0.566
24-hour	0.853	1.000	1.000	0.857	0.762	0.827
24-hour biosimilar	0.324	0.754	0.993	0.310	0.369	0.396
36-hour	0.516	0.954	0.994	0.512	0.449	0.629
42-hour	0.316	0.863	0.944	0.307	0.272	0.577
Type 2	0.290	0.856	0.934	0.283	0.220	0.477
	B. Switching behavior					
Ever switching	0.301	0.996	1.000	0.283	0.348	0.650
Switches (Share)		0.213	0.529	0.213	0.029	0.016
Switching × 24-hour biosimilar	0.122	0.574	0.952	0.104	0.192	0.233
Switching × 36-hour	0.144	0.910	0.900	0.129	0.141	0.461
Switching × 42-hour	0.068	0.811	0.830	0.051	0.081	0.466
Switching × Type 2	0.052	0.796	0.742	0.037	0.056	0.350

Note: Physician-level summary statistics. Patients, visits and prescriptions are restricted to patients in the sample of interest (Table 1.11). Patients, visits, CGM users and CGM visits correspond to the sample median.

Table 1.13: Predicted vs observed CGM adoption

		Observed	
		No	Yes
Predicted	No	0.8243	0.2463
	Yes	0.1757	0.7537

Note: Table 1.13 presents the accuracy of CGM adoption prediction for patients adopting the technology from January 2018. Adoption is predicted from individual demographics and the evolution of strip consumption over time for patients using the technology. Youden's index is used as a threshold to classify observations. The risks of Type I and Type II errors are 17.6% and 24.6%, respectively.

Table 1.14: Switching costs, CGM users vs non-users

	A. All		B. Specialists	
	coef.	s.e.	coef.	s.e.
α	-0.186	(0.039)	-0.232	(0.043)
c_1: Incumbent	-2.010	(0.039)	-2.195	(0.045)
c_2: Users	0.013	(0.053)	-0.091	(0.058)
c_3: Using CGM	0.024	(0.112)	0.082	(0.119)
N	621		414	

Note: Robust standard errors. The first column includes prescriptions from GPs, diabetes specialists and diabetes specialists working in the hospital. The second column focuses on specialists. Both specifications include a quadratic time trend.

Table 1.15: Information set and non-user switching behavior

	A. All non-users		B. Eligible		C. Non-eligible	
	coef.	s.e.	coef.	s.e.	coef.	s.e.
$\mathbf{1}(F_{kt} > 0)$	0.548	(0.036)	0.243	(0.017)	0.306	(0.025)
F_{kt}	0.159	(0.016)	0.059	(0.009)	0.096	(0.011)
$(F_{kt})^2$	-0.003	(0.000)	-0.001	(0.000)	-0.002	(0.000)
F_{kt}^{CGM}	-0.081	(0.027)	-0.012	(0.015)	-0.061	(0.016)
$(F_{kt}^{CGM})^2$	0.001	(0.001)	-0.001	(0.001)	0.001	(0.001)
Obs.	7,568		7,434		7,497	
$\overline{Y}$	0.932		0.428		0.517	
$\overline{F}$	3.709		3.728		3.723	
$\overline{F}_{CGM}$	2.288		2.292		2.292	

Note: The information set is measured in $'0$ of visits. Standard errors clustered at the physician level. I restrict the sample to individual diabetes specialists and focus on the 42-hour product.

Table 1.16: Accurate prediction, per year and patient group

Cluster	2015	2016	2017	2018	2019	2020	2021
1	0.82	0.73	0.62	0.50	0.52	0.54	0.54
2	0.79	0.69	0.55	0.44	0.47	0.47	0.45
3	0.72	0.64	0.52	0.43	0.49	0.53	0.52
4	0.77	0.68	0.47	0.43	0.39	0.42	0.42
5	0.72	0.65	0.48	0.41	0.40	0.40	0.41
6	0.56	0.51	0.38	0.34	0.33	0.31	0.32
7	0.61	0.56	0.38	0.34	0.33	0.32	0.31

Note: This table compares the product prescribed in the actual data j^{obs} to $\hat{j} = \arg\max u_{ikjt}(p_{jt}, d_j, a_{it}|\mathcal{I}_{kt})$. The idiosyncratic shock is excluded from the prediction. The accurate prediction rates range from 31% to 82%.

Table 1.17: Patient selection

	(1) All, incl:	(2) Final Sample	(3) Infrequent	(4) New Patients	(5) No diabetes specialist	(6) Old Age≥ 75	(7) Pump	(8) Rest
N ('000)	1,423.0	333.4	54.3	335.0	91.9	408.7	48.3	151.5
Age	63	57	55	52	61	82	44	58
Female	0.483	0.434	0.491	0.474	0.425	0.558	0.560	0.420
Low income	0.103	0.126	0.156	0.162	0.122	0.018	0.105	0.135
Type 1	0.158	0.283	0.123	0.093	0.163	0.080	0.731	0.060
Death	0.003	0.002	0.005	0.001	0.001	0.007	0.000	0.002
Residential area								
Deprivation index	0.334	0.262	0.296	0.295	0.627	0.331	0.201	0.475
Population ('000)	38.3	41.0	43.8	42.2	31.8	34.6	32.2	37.5
Chronic conditions								
Hypertension	0.701	0.692	0.621	0.521	0.745	0.897	0.405	0.657
Hypercholesterolemia	0.550	0.619	0.453	0.415	0.649	0.616	0.370	0.527
Analgesics	0.429	0.412	0.459	0.335	0.419	0.544	0.272	0.389
Cardiovascular	0.358	0.318	0.440	0.205	0.249	0.577	0.171	0.264
Obesity	0.250	0.324	0.401	0.206	0.217	0.220	0.263	0.216
Anxiolytics	0.193	0.174	0.219	0.154	0.171	0.251	0.127	0.185
Antidepressant	0.178	0.168	0.180	0.139	0.145	0.230	0.161	0.163
Respiratory	0.158	0.148	0.227	0.128	0.126	0.196	0.109	0.150
Hypnotics	0.122	0.108	0.144	0.090	0.108	0.166	0.069	0.115
Cancer	0.137	0.097	0.260	0.089	0.087	0.217	0.056	0.125
Neuroleptics	0.047	0.043	0.060	0.044	0.041	0.055	0.019	0.055
Dialyse	0.013	0.016	0.048	0.005	0.004	0.016	0.006	0.007
Long-term care	0.842	0.927	0.691	0.746	0.930	0.845	0.939	0.828
Short-acting	0.556	0.744	0.610	0.512	0.474	0.491	0.982	0.313
Specialists	0.720	1.000	1.000	1.000	0.022	0.606	0.895	0.058
Prescriptions (Av.)	16	27	7	9	29	16	8	9
Specialist prescriptions (Av.)	3	7	3	3	0	2	4	0
CGM Users	0.226	0.449	0.023	0.214	0.125	0.081	0.930	0.065
CGM Temporary	0.091	0.073	0.364	0.128	0.136	0.138	0.018	0.188
Pump Users	0.051	0.067	0.000	0.000	0.000	0.004	1.000	0.000
Insulin switch	0.381	0.676	0.200	0.307	0.329	0.321	0.445	0.129
Nb switches	1.673	1.848	1.282	1.523	1.516	1.609	1.673	1.332

Note: The Deprivation index is computed based on 2015 measures of unemployment, blue-collar workers, high school graduates shares and the median income by consumption unit at the city level from the national statistic institute (INSEE). It is centered around zero, goes from -6.1 to 10.3, and the variance is 2.72. Negative values stand for more favorable areas. In 2015, the median individual in France lived in a 9,423 inhabitants city, and the deprivation index is around 0.116. The number of insulin switches is computed on the sample of patients who switched at least once. 'Infrequent' includes patients using insulin spontaneously over the sample period or who stopped before April 2016. 'New patients' have their first insulin prescription after March 2016. 'No diabetes specialist' refers to patients who have not seen a specialist already active in 2016. 'Old' patients were 75 or more in 2015. 'Pump' includes patients relying exclusively on insulin pumps. 'Rest' consists of the remaining patients prescribed insulin by new physicians or physicians who are not actively changing treatments.

1.9 Data construction

1.9.1 Sample selection

I rely on exhaustive micro-level data for France from the *Système National des Données de Santé (SNDS)*. The data on the French population is exhaustive, thanks to the single-payer system. The Social Security system covers more than half of prescription drug expenditures and medical devices once the Health Technology Agency has enacted the coverage. The mandatory health insurance system fully covers patients with diabetes for their expenditures as part of the Long-term care disease program.[48] I build the sample of patients as follows.

First, I retrieve all the long-acting insulin reimbursement flows from 2015 to 2021. I extract the patient, prescriber, product IDs, and prescription date for each flow.[49] Note that the prescriber ID corresponds to the hospital ID for physicians working at the hospital. On the patient side, I focus on patients who were already familiar with insulin injections and glucose measurement before April 2016. I impose the following restrictions among patients: (i) The first long-acting insulin prescription was recorded before April 1st, 2016 and the last prescription was written after January 1st, 2016. (ii) The time between the first and last prescription was over three years. (iii) The patient received more than one prescription per year between his first and last prescription. (iv) I restrict my attention to patients not using an insulin pump or whose first pump reimbursement happened after January 1st, 2016.[50] (v) The patient was an adult, 74 or younger, in 2015. On the physician side, I consider (i) The diabetes specialists working in or out of the hospital. (ii) The prescribers with a first prescription to *any patient* before January 2016. (iii) The practitioners who wrote at least 24 prescriptions and switched a treatment at least once for *any patients in my sample* between 2016 and 2021. Table 1.17 compares the characteristics of patients in the final sample to the initial dataset. Focusing on the set of incumbent insulin patients, Table 1.12 provides summary statistics about the physicians involved in their therapy.

Most of the descriptive analysis relies on this sample. Note that (i) The first intention treatment choice for new insulin patients is used to estimate switching costs separately from

[48] The drug price still enters the care provider's decision through financial incentive schemes and guidelines.

[49] I focus on the following Anatomical Therapeutic Chemical (ATC) classes: A10AC, A10AD, A10AE. If two insulins in these classes are prescribed on the same day, I consider either (i) the new molecules when a new product is prescribed together with an old product or (ii) the 24-hour product when it is prescribed at the same time as insulin mixes. The remaining cases are dropped.

[50] The data also covers insulin pump coverage. Patients using a pump to stabilize their glucose level are prescribed long-acting insulin in case of a breakdown. They are not using it daily. When a patient starts using an insulin pump, I remove the long-acting insulin prescriptions written while using the pump.

learning (Section 1.10.3). (ii) The structural model is estimated on a random sample of 150 diabetes specialists working outside the hospital.

1.9.2 Risk of mismeasurement in CGM adoption

This section documents the risk of mismeasurement in digital device adoption when relying on the claims data. It considers sequentially the risk for patients eligible for coverage and non-eligible for CGM coverage.

For individuals eligible for the technology, purchases of sensors directly from the device manufacturer are not reimbursed by the health insurance scheme. It is fully covered in the case of a prescription. These patients have no incentive to buy the technology outside the scope of my data. The coverage decision comes 2.5 years after the device was available in the EU, such that the adoption date imputed from the claims data is subject to left censoring. This concern is mitigated by the low sales volumes in 2015 for the most popular technology (40k sensors according to the HTA evaluation) and by the timing of adoption in the population according to claims data. The share of patients with a first sensor reimbursement in the first three months following the coverage decision remains limited (Figure 1.2a). In the descriptive analysis, I remove patients adopting CGM between June and August 2017 to account for the risk of mismeasurement in their adoption date. Individuals dropping out from using digital devices is an important issue when studying the impact of wearable technologies (Patel et al. (2015)). Technological features of the current device allow for overcoming this challenge. CGM relies on disposable sensors that must be replaced every 14 days. Patients not renewing their prescriptions are assumed to drop out from continuous monitoring. In practice, such cases are rare due to the comfort brought by the device to the patient.

Among patients treated with long-acting insulin, patients ineligible for coverage represent 32% of individuals (Table 1.11). In this case, measurement error in device adoption is possible. First, Guerci et al. (2023) suggest that patients outside of the eligibility criteria got access to the technology as no prior authorization was required. Second, the alternative glucose measurement system relies on disposable strips such that the number of strips reimbursed represents a good proxy for glucose testing intensity. If individuals were to adopt the technology outside the insurance system, strip reimbursement for these patients would decrease as measurement systems are substitutes to each other. I study the evolution of strip reimbursements for CGM users and nonusers over time in Figure 1.26. Figure 1.26a plots the average daily number of

strips reimbursed before and after CGM adoption. Strip consumption decreases by 2.5 strips on average after adopting the technology.[51] The average for nonusers stays relatively constant despite fluctuations across quarters. This suggests that the consumption of nonuser glucose testing strips remains unaffected by the introduction of CGM. These descriptive facts do not rule out completely the adoption of the technology outside of the insurance scheme but suggest that it is uncommon.

In the structural analysis, the timing of adoption matters as it determines the information available to the physician when prescribing insulin. Patients whose first CGM prescription happens within the seven months following the coverage decision may have purchased the technology out-of-pocket before. As a result, their actual adoption date might not align with the first prescription recorded in the claims data. To accurately estimate the adoption date for these 'Early users', I leverage the substitutability between CGM and glucose test strips as strip consumption decreases once a patient adopts a CGM (Figure 1.26a). More precisely, I split the sample of CGM users into two distinct groups: (i) 'uncensored' eligible patients whose first CGM reimbursement occurs from 2018 onwards and (ii) 'censored' patients whose first reimbursement occurs before January 2018. I estimate the correlation between strip consumption and CGM adoption in the 'uncensored' population. I extrapolate these findings to the 'censored' population to infer their 'true' date of adoption. In particular, I estimate the following model, linking the number of test strips reimbursed to patient i in quarter q and CGM use:

$$CGM_{iq} = \alpha_0 strips_{iq} + \gamma \mathbf{1}(strips_{i,t+1})\mathbf{1}(strips_{i,t+2})strips_{i,2015} + \beta X_i + \varepsilon_{iq} \tag{1.28}$$

where $strips_{iq}$ measures the average number of strips reimbursed per day per quarter, and X_i includes diabetes type, gender and age. I use the estimates out-of-sample to predict the quarter of adoption in the 'censored' population, relying on Youden's index to classify observations based on their predicted probability. Table 1.13 reports the risks of Type I and Type II errors in the 'uncensored' sample, respectively 17.6% and 24.6%. I consider that a patient in the 'censored' sample adopted the technology at t if the model predicts adoption for three quarters in a row. Figure 1.27 plots the cumulative distributions of observed and predicted adoption quarters for early users. The CGM adoption quarter differs for approximately 18% of the patients.

[51] Patients keep consuming strips after CGM adoption to perform a regular test in case of an adverse event.

1.10 Further descriptive evidence

1.10.1 Substitution towards insulin pump

Patients can use an insulin pump instead of the short-acting and long-acting insulin to stabilise their glucose levels. These patients rely exclusively on short-acting insulin and only get prescribed long-acting insulin in case of a pump breakdown. The long-acting insulins prescribed to pump users are excluded from the prescriptions physicians are learning from. Significant improvements in the pump technology have occurred over the past decade. The pump system is not in the physician's choice set in the main specification. This choice is motivated by two facts. First, the coverage for insulin pumps remains restricted in France. Second, while patients in my sample are increasingly relying on a pump system, there is no change in the pattern of pump adoption after the coverage of CGM (Figure 1.25), and the total number of patients remains limited (70k). There exists a discrepancy with the share of pump users displayed in Table 1.11, primarily because the table excludes patients relying on a pump system for the entire period and focuses on patients with long-acting insulin prescriptions, hence does not represent the share of pump users in the overall patient's population.

1.10.2 Detailing as an alternative source of learning

In this project, pharmaceutical detailing may matter as it represents an alternative source of learning about new products for patients and physicians (Grennan et al. (2024)). Direct-to-consumer advertising is forbidden for prescription drugs in France, and detailing to physicians is allowed but subject to transparency rules similar to the US Sunshine Act. Two caveats prevent me from precisely accounting for detailing to physicians. First, unlike the Sunshine Act data, the French publicly available data does not record the product mentioned during the meeting. The probability that the interaction mentioned insulins depends on the pharmaceutical companies' portfolio and the physician's medical speciality. Detailing to diabetes specialists is more likely to mention new insulins than interactions with GPs. The latter is even more uncertain as the company portfolio is diverse. Long-acting insulins account for around 8% of Sanofi's European sales in 2023. On the other hand, diabetes treatment represents 88% of Novo Nordisk sales in 2023, so its detailing strategy is likely to trigger diabetes. Second, the claims data cannot be linked to external sources, including the *Transparence Santé* registry. In Figure 1.29, I provide summary statistics about insulin manufacturer's detailing behavior towards diabetes specialists

and GPs separately.

1.10.3 Switching cost estimation

I consider separately the demand for the 24-hour product from physicians of medical speciality, m, for patients in group, $I \in \{N, E^A, E^{\bar{A}}\}$, among new patients, N, existing patients adopting, E^A, and not adopting CGM, $E^{\bar{A}}$. Denoting s_{jq}^{mI} the prescription share of Lantus (L) and its biosimilar (B) indexed by $j \in \{L, B\}$ and at time q, I estimate switching costs relying on the following model

$$log(s_{Bq}^{mI}) - log(s_{Lq}^{mI}) = \Delta\delta_q^m + \alpha\Delta p_q^m + c_1\mathbf{1}(I \in \{E^{\bar{A}}, E^A\}) + (c_2 + c_3 CGMv_q^m)\mathbf{1}(I = E^A) + \Delta\xi_q \tag{1.29}$$

$log(s_{Bq}^{mI}) - log(s_{Lq}^{mI})$ proxies the difference in mean utility between the biosimilar, B, and the branded version, Lantus, L. $\Delta\delta_q^m$ is a time trend which accounts for information frictions. Δp_q^m is the difference in prices.[52] $CGMv_q^m$ is the share of prescriptions to patients wearing a CGM. c_1, c_2, and c_3 arise only in the biosimilar demand for incumbent patients and proxy switching costs. The model is estimated by OLS, and the results are displayed in Table 1.14. $\widehat{c}_1 < 0$ suggesting positive switching costs, yet no heterogeneity across glucose measurement systems as $\widehat{c}_2$ and $\widehat{c}_3$ are not significant. From these estimates, the willingness to pay to stick to the treatment patients are familiar with lies around ~ 9.47€ per month.

1.10.4 Physician level-learning: qualitative effect of CGM information

For each physician k and quarter t, I count the number of switches to the 42-hour product for patients without a CGM. Physicians' information set, denoted F_{kq}, is approximated using the number of appointments up to $t-1$ where the patient already used the 42-hour product at the beginning of the visit. The correlation between F_{kq} and the number of switches to the 42-hour product in quarter t, Y_{kq}, is documented using a Poisson model:

$$E(Y_{kq}|X) = \exp\left(\alpha + \lambda_1\mathbf{1}(F_{kq} > 0) + f(F_{kq}) + \gamma X_k + \delta_q\right) \tag{1.30}$$

[52] Given the pricing scheme in France, the price coefficient is difficult to identify separately from the time trend. Here, it is identified by exploiting the introduction of a financial scheme for diabetes specialists working in the hospital, rewarding the biosimilar prescription over the branded 24-hour product. This variation is assumed to be uncorrelated with the error term.

f is a quadratic function of F_{kq}. $\mathbf{1}(F_{kq} > 0)$ captures the extensive margin. The model is estimated from physician-quarter combinations with more than 11 visits. Table 1.15 presents the results. Going from 10 experience feedbacks to 20 increases the occurrence of switches to the 42-hour product by 16.2%. Digital technology feedback seems to dampen the spillovers. Specifications B and C consider the correlation between the information set and switching eligible nonusers and non-eligible patients separately. Eligible patients are more similar to CGM users in terms of observable and unobservable characteristics than non-eligible. For eligible individuals, receiving feedback from CGM against the traditional measurement system has no significant impact on the magnitude of the spillover effect. For non-eligible individuals, the spillover effect is significantly weaker if the feedback is received from a patient using a CGM. If physicians extrapolate based on patients ' similarities, this effect may be driven by patient heterogeneity (Alsan et al. (2024)).

1.11 Multi-product bargaining

The regulator and the pharmaceutical company, f, are assumed to bargain over the full set of insulins offered by the manufacturer, $j \in \mathcal{J}_f$. The objective of the firm is to maximize its profits:

$$\pi_{ft}(\mathbf{p}_t) = \sum_{\forall j \in \mathcal{J}_f} (p_{jt} - c_{jt}) q_{jt}(\mathbf{p}_t) \tag{1.31}$$

The objective of the regulator when bargaining with firm f is to maximise consumer surplus:

$$\Delta_{ft} CS(\mathbf{p}_t) = \sum_{\forall i,k} \frac{1}{\lambda} \ln\Big(\sum_{\forall j \in \mathcal{J}} \exp(u_{ikjt}(p_{jt}, d_j, a_{it}|\mathcal{I}_{kt})\Big) - \sum_{\forall i,k} \frac{1}{\lambda} \ln\Big(\sum_{\forall j' \notin \mathcal{J}_f} \exp(u_{ikj't}(p_{j't}, d_{j'}, a_{it}|\mathcal{I}_{kt})\Big) \tag{1.32}$$

Where we allow λ to be different from α. The equilibrium prices maximize the Nash product:

$$\max_{\mathbf{p}_{ft}} [\pi_{ft}(\mathbf{p}_t)]^{b_{ft}} [\Delta_{ft} CS(\mathbf{p}_t)]^{1-b_{ft}} \tag{1.33}$$

Let's consider the case where the firm bargain over two products, indexed by j and j'. The FOC with respect to the price of product j, p_{jt}, is given by:

$$b_{ft} \frac{\partial \pi_{ft}(\mathbf{p}_t)/\partial p_{jt}}{\pi_{ft}(\mathbf{p}_t)} + (1 - b_{ft}) \frac{\partial \Delta_{ft} CS(\mathbf{p}_t)/\partial p_{jt}}{\Delta_{ft} CS(\mathbf{p}_t)} = 0 \tag{1.34}$$

where

$$\frac{\partial \pi_{ft}(\mathbf{p}_t)}{\partial p_{jt}} = q_{jt}(\mathbf{p}_t) + (p_{jt} - c_{jt})\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}} + (p_{j't} - c_{j't})\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}}$$

and

$$\frac{\partial \Delta_{ft} CS(\mathbf{p}_t)/\partial p_{jt}}{\Delta_{ft} CS(\mathbf{p}_t)} = \frac{-\alpha q_{jt}(\mathbf{p}_t)}{\Delta_{ft} CS(\mathbf{p}_t)} = \frac{\sum -\alpha s_{ikjv}(\mathbf{p}_t)}{\Delta_{ft} CS(\mathbf{p}_t)}$$

The scaling parameter, λ does not enter the FOC, hence has not impact on the equilibrium. Denoting $h_{jt} = \frac{\partial \Delta_{ft} CS(\mathbf{p}_t)/\partial p_{jt}}{\Delta_{ft} CS(\mathbf{p}_t)}$ and $\beta_{ft} = \frac{1-b_{ft}}{b_{ft}}$, the first order condition with respect to p_{jt} becomes

$$\begin{aligned}
&\frac{\partial \pi_{ft}(\mathbf{p}_t)}{\partial p_{jjt}} + \beta_{ft} h_{jt} \pi_{ft}(\mathbf{p}_t) = 0 \\
&q_{jt}(\mathbf{p}_t) + \Big(p_{jt} - c_{jt}\Big)\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}} + \Big(p_{j't} - c_{j't}\Big)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}} + \beta_{ft} h_{jt}\Bigg((p_{jt} - c_{jt})q_{jt}(\mathbf{p}_t) + (p_{j't} - c_{j't})q_{j't}(\mathbf{p}_t)\Bigg) = 0 \\
&(p_{jt} - c_{jt})\Bigg(\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}} + \beta_{ft} h_{jt} q_{jt}(\mathbf{p}_t)\Bigg) + (p_{j't} - c_{j't})\Bigg(\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}} + \beta_{ft} h_{jt} q_{j't}(\mathbf{p}_t)\Bigg) + q_{jt}(\mathbf{p}_t) = 0
\end{aligned} \tag{1.35}$$

By symmetry, the FOC with respect to $p_{j't}$ yields

$$\begin{aligned}
&(p_{j't} - c_{j't})\Bigg(\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} + \beta_{ft} h_{j't} q_{j't}(\mathbf{p}_t)\Bigg) = -\Bigg[(p_{jt} - c_{jt})\Bigg(\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}} + \beta_{ft} h_{j't} q_{jt}(\mathbf{p}_t)\Bigg) + q_{j't}(\mathbf{p}_t)\Bigg] \\
&(p_{j't} - c_{j't}) = -\Bigg[(p_{jt} - c_{jt})\Bigg(\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}} + \beta_{ft} h_{j't} q_{jt}(\mathbf{p}_t)\Bigg) + q_{j't}(\mathbf{p}_t)\Bigg]\Bigg[\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} + \beta_{ft} h_{j't} q_{j't}(\mathbf{p}_t)\Bigg]^{-1}
\end{aligned} \tag{1.36}$$

Plugging Equation 1.36 into Equation 1.35,

$$\begin{aligned}
(p_{jt} - c_{jt})\Bigg[\Big(\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}} + \beta_{ft} h_{jt} q_{jt}(\mathbf{p}_t)\Big) - \Big(\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}} + \beta_{ft} h_{j't} q_{jt}(\mathbf{p}_t)\Big)\Big(\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} + \beta_{ft} h_{j't} q_{j't}(\mathbf{p}_t)\Big)^{-1}\Big(\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}} + \beta_{ft} h_{jt} q_{j't}(\mathbf{p}_t)\Big)\Bigg] = \\
-q_{jt}(\mathbf{p}_t) + q_{j't}(\mathbf{p}_t)\Big(\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} + \beta_{ft} h_{j't} q_{j't}(\mathbf{p}_t)\Big)^{-1}\Big(\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}} + \beta_{ft} h_{jt} q_{j't}(\mathbf{p}_t)\Big)
\end{aligned} \tag{1.37}$$

such that

$$\begin{aligned}
(p_{jt} - c_{jt})\Bigg[\Big(\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}} + \beta_{ft} h_{jt} q_{jt}(\mathbf{p}_t)\Big)\Big(\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} + \beta_{ft} h_{j't} q_{j't}(\mathbf{p}_t)\Big) - \Big(\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}} + \beta_{ft} h_{j't} q_{jt}(\mathbf{p}_t)\Big)\Big(\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}} + \beta_{ft} h_{jt} q_{j't}(\mathbf{p}_t)\Big)\Bigg] = \\
-q_{jt}(\mathbf{p}_t)\Big(\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} + \beta_{ft} h_{j't} q_{j't}(\mathbf{p}_t)\Big) + q_{j't}(\mathbf{p}_t)\Big(\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}} + \beta_{ft} h_{jt} q_{j't}(\mathbf{p}_t)\Big)
\end{aligned} \tag{1.38}$$

Given $h_{jt} = \frac{-\alpha q_{jt}(\mathbf{p}_t)}{\Delta_{ft} CS(p)}$,

$$\beta_{ft} h_{j't} q_{jt}(\mathbf{p}_t) = \beta_{ft} h_{jt} q_{j't}(\mathbf{p}_t)$$
$$\beta_{ft} h_{j't} q_{j't}(\mathbf{p}_t) q_{jt}(\mathbf{p}_t) = \beta_{ft} h_{jt} q_{j't}(\mathbf{p}_t) q_{j't}(\mathbf{p}_t)$$

the first order condition becomes,

$$\begin{aligned}(p_{jt} - c_{jt})\Bigg[\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}}\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} + \beta_{ft} h_{jt}\left(q_{jt}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} - q_{j't}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}}\right) - \\ \frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}}\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}} + \beta_{ft} h_{j't}\left(q_{j't}(\mathbf{p}_t)\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}} - q_{jt}(\mathbf{p}_t)\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}}\right)\Bigg] = \\ -\left(q_{jt}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} - q_{j't}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}}\right)\end{aligned} \tag{1.39}$$

such that

$$c_{jt} = p_{jt} + \left[\beta_{ft} h_{jt} + \frac{\beta_{ft} h_{j't}\left(q_{j't}(\mathbf{p}_t)\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}} - q_{jt}(\mathbf{p}_t)\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}}\right) + \left(\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}}\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} - \frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}}\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}}\right)}{\left(q_{jt}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} - q_{j't}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}}\right)}\right]^{-1} \tag{1.40}$$

$$c_{jt} = p_{jt} + \left[\beta_{ft} h_{jt} + \frac{\partial q_{jt}(\mathbf{p}_t)/\partial p_{jt}}{q_{jt}(\mathbf{p}_t)} + \left(\beta_{ft} h_{j't} + \frac{\partial q_{j't}(\mathbf{p}_t)/\partial p_{jt}}{q_{jt}(\mathbf{p}_t)}\right)\frac{\left(q_{j't}(\mathbf{p}_t)\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}} - q_{jt}(\mathbf{p}_t)\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}}\right)}{\left(q_{jt}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{j't}} - q_{j't}(\mathbf{p}_t)\frac{\partial q_{j't}(\mathbf{p}_t)}{\partial p_{jt}}\right)}\right]^{-1} \tag{1.41}$$

For single-product firms, the first-order condition boils down to

$$c_{jt} = p_{jt} + \left[\beta_{ft} h_{jt} + \frac{\partial q_{jt}(\mathbf{p}_t)/\partial p_{jt}}{q_{jt}(\mathbf{p}_t)}\right]^{-1} \tag{1.42}$$

Note that, due to the Type I extreme value error term,

$$\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{jt}} = \sum_{\forall i,k}\sum_{\forall v \in t} -\alpha s_{ikjv}(\mathbf{p}_t)(1 - s_{ikjv}(\mathbf{p}_t)) \tag{1.43}$$

and, for j and j', $j' \neq j$

$$\frac{\partial q_{jt}(\mathbf{p}_t)}{\partial p_{j't}} = \sum_{\forall i,k}\sum_{\forall v \in t} \alpha s_{ikjv}(\mathbf{p}_t) s_{ikj'v}(\mathbf{p}_t) \tag{1.44}$$

The remaining unobservable are marginal costs, c_{jt}, and bargaining weights, b_{ft} .

1.12 Counterfactual scenario: sensor reallocation

The counterfactual in Section 1.6.2 suggests that information spillovers from CGM users to nonusers are limited. The consumer surplus generated by the technology mostly accrued to patients using the device. As Table 1.1 suggests that adoption is heterogeneous across physicians, the lack of spillovers could be driven by little variation in CGM adoption within the set of patients followed by a given physician. Considering the extreme case where all patients within a cluster at a particular practice are either adopting or non-adopting, nonusers cannot benefit from the information generated by CGM users. To analyze whether this phenomenon drives the lack of spillovers, I compute the equilibrium shares and the consumer surplus generated by spillovers under an alternative allocation of CGM, which removes the variation across physicians. To that end, first, I predict patient device adoption in my sample of physicians using a model similar to the one from Table 1.1. I compute the predicted probability of adoption, net of the physician fixed effect for each patient, $\hat{p}_i$. Two patients with similar characteristics have the same probability of adopting the device. I randomly select patients to adopt the technology, fixing the number of users to the number observed in the data and using the predicted probability of adoption, $\hat{p}_i$, as weights. Adoption dates are allocated randomly. I compute consumer welfare under this alternative allocation and present the results (against the allocation observed in the data) in Figure 1.34.

Chapter 2

Strategic Tier Design in Health Insurance: The Case of Medicare Part D

Abstract

We study the role of tier design in Medicare Part D. In the period 2013-2017, plans expanded the number of tiers in their formularies from three/four to five and systematically shifted generics to higher tiers subject to higher cost sharing. The systematic tier upgrading caused significant increases in the out-of-pocket costs, up to 6 times for some generics. This resulted in additional average per-enrollee spending on generics of $76 in 2017, totalling $1.5 billion for the Part D population, and increased mortality by 5.4% due to reduced utilization of generics with documented mortality benefits.

2.1 Introduction

In industries characterized by complex products or services, consumers may be unable to process every detail, allowing firms to leverage the complexity of their offers to increase profits (Ellison and Ellison (2009); Carlin (2009); Armstrong and Chen (2009)). Medicare Part D contracts are notoriously complex: the price is a combination of a premium, a deductible, and non-linear coverage rates that dynamically depend on utilization throughout the year (Dalton et al. (2019)). This makes comparisons difficult: as premiums are clear and certain, while future consumption is not, enrollees tend to base their plan choices mostly on premiums rather than on cost sharing or drug coverage (Handel and Kolstad (2015); Abaluck and Gruber (2011, 2016); Decarolis et al. (2020)). Whereas inertia in enrollees' plan choice led to a steady rise of premiums and triggered stricter regulations around premium setting (Marzilli Ericson (2014); Decarolis (2015); Ho et al. (2017); Fleitas (2020)), the structure of formularies has remained relatively unregulated.

In this paper, we study strategic tier design in Medicare Part D. Plans are required to cover at least two drugs in the most commonly prescribed categories and all drugs that treat certain conditions. However, insurers are free to choose both the type (copayment or coinsurance) and the level of cost sharing, which usually increases in the tier, as long as the plan is actuarially equivalent to a standard plan. Insurers use the tier structure to steer consumption to cheaper types of medical care. First, since cheaper therapeutically equivalent substitutes (mostly generics) are provided at lower cost sharing, the tier structure increases demand elasticity to price and insurers' bargaining power vis-à-vis drug manufacturers, driving list prices down (Duggan and Scott Morton (2010); Olssen and Demirer (2021)). Second, plans that bundle pharmaceutical and medical benefits internalize the externality of drug utilization on inpatient hospitalization, thus offering more generous drug coverage (Lavetti and Simon (2018); Starc and Town (2020)).

We use publicly available data on formularies, enrolment, and drug utilization for Medicare Part D stand-alone prescription drug plans between 2013 and 2017 to investigate tier design by insurers and their impact on enrollees' out-of-pocket costs, demand, and spending for generics. We document that, over this period, plans made larger use of five-tiered formularies, which include two generic tiers, two brand tiers, and a specialty tier. Over time, plans increasingly placed generics in higher tiers (Avalere (2018); Fix and Whorley (2022)), even those usually reserved for branded drugs and subject to higher cost sharing—effectively making the cheapest available treatments more expensive for millions of enrollees. This shrunk the size of the first

tier (with the lowest cost sharing), traditionally devoted to generics, which in 2017 included three times fewer drugs than in 2013.

To quantify the effect of tier shifting on out-of-pocket costs, we exploit variation across generic drugs that were moved to higher tiers. To address the endogeneity of tier placement, we leverage the rules on actuarial equivalence and build an instrument that links the probability of tier shifting for each drug to the distance between the effective cost-sharing in the plan and 25%, the rule for the standard plan. The systematic tier upgrading of generics caused an increase in the out-of-pocket costs for the average monthly supply of 6 times for high value generics and of 0.4 times for other generics. This is not due to a handful of high-utilization drugs or a few high-enrolment plans, but affected expensive and cheap generics alike and was so widespread that, even if a sophisticated enrollee could perfectly adjust drug choices (for example choosing therapeutic substitutes in lower tiers) or switch plan, they would still face higher out-of-pocket costs in 2017 than in 2013. We exclude that this increase in out-of-pocket costs was due to a generalized increase in drug prices.

To assess some of the consequences of higher out-of-pocket costs due to tier upgrading of generics, we rely on drug utilization data and estimate a demand model linking purchases of generics to their tier placement. To overcome the endogeneity of tier placement in demand estimation, we build Hausman-type instruments (Hausman (1996); Nevo (2001)). We instrument the tier of a specific drug in a region with the average tier of the same drug in other regions. Controlling for drug, region, and year fixed effects, and excluding shared formularies from the computation, the validity of this instrument rests on the absence of residual aggregate demand shocks that could lead to nation-wide tier shifting of a drug across plans.

We estimate low price elasticities in line with those by Einav et al. (2018), but still large enough to reduce consumption compared to a counterfactual scenario in which generics had remained in the same tier as in 2013. Our results illustrate that the average enrollee spent \$76 more on generic drugs in 2017 due to tier upgrading. Enrollees who instead adhered to therapy in spite of the higher out-of-pocket costs spent \$100 more. Collectively, these increases correspond to an additional \$1.5 billions in out-of-pocket spending for the Medicare Part D population in 2017.

Notably, the generics in our sample include high value medications with documented mortality benefits: if utilization dropped due to higher out-of-pocket costs, enrollees may have suffered adverse health outcomes. A back-of-the-envelope calculation in the spirit of Chandra

et al. (2021) shows that reduced utilization due to tier upgrading for drugs in classes with documented mortality benefits resulted in increases in mortality by 5.4%. We also document an additional welfare-decreasing effect of tier upgrading: contrary to what insurance theory deems optimal (Besley (1988); Feldstein (1973); Einav et al. (2018)), the generics moved to higher tiers were those with lower demand elasticity to price, hence those less prone to moral hazard. By distorting the optimal cost sharing balancing risk protection and moral hazard, tier upgrading may have thus negatively affected welfare beyond reductions in utilization and increases in spending.

We interpret our results as lower bounds for the effects of tier shifting on expenditures and welfare, as our estimates do not capture the full substitution across drugs, health outcomes, and public spending. First, the combination of higher cost sharing for generics and the increasing role of rebates may favor brand-name medications (Dafny et al. (2017); Dubois and Lasio (2018); Dusetzina et al. (2019)). In turn, lower use of generics may affect long-run entry decisions of generic manufacturers, leading to even fewer alternatives, lower competition, and further risk of shortages in the future (Berndt et al. (2017)). Higher cost sharing may also result in higher public spending. For enrollees with the low-income subsidy, Medicare fully covers the out-of-pocket costs. Also, enrollees may reach the catastrophic coverage threshold at a faster pace ($4,950 in 2017), after which Medicare covers 80% of the expenses. Finally, lower therapy adherence may result in inefficient substitution to other types of medical care (Chandra et al. (2010); Starc and Town (2020)), with the potential for increased hospitalization and higher costs for society.

The rest of the paper is organized as follows. Section 2.2 describes the empirical setting and data. Section 2.3 provides evidence of the systematic tier upgrading of generics. The impact of tier upgrading on out-of-pocket costs is studied in Section 2.4 and on demand and total spending in Section 2.5. Section 2.6 discusses some welfare implications of tier upgrading. Section 2.7 concludes.

2.2 Data

We use three sources of data from Medicare Part D, available from the Centers for Medicare and Medicaid Services (CMS). First, we rely on formulary data from 2013-2017 to retrieve information on the characteristics of the formulary for each plan-year: the list of drugs identified

by their unique National Drug Code (NDC), the tier they belong to, the cost sharing rules, and plan-level list prices (for 2013 and 2017 only). We use the NDC to identify the Anatomical Therapeutic Chemical (ATC) codes from the WHO for each drug, as well as the drug type (branded/generic). Second, we use the Prescriber Event file to recover quantities (in 30-day supply) and the total cost for each drug. Finally, we refer to enrolment data to compute plan-level market shares.

Following Einav et al. (2018), a drug is defined by its non-proprietary name and whether it is branded or generic. If a drug appears in several tiers, we consider the lowest; if it appears in different dosages or forms in the same tier, we consider the least expensive version. A drug class is defined at the level 4 of the WHO ATC classification (ATC4), our measure of the product market. We follow the literature and exclude from the analysis drug classes with less than 100,000 claims over the period 2013-2017 for the main 34 geographical markets. Quantities are defined in 30-day supply units. We combine plan-level formulary data and enrolment shares to recover the average tier of a drug in a given market for a given year. We deflate list prices and OOP costs using the US CPI with 2013 as the base year.

We focus on stand-alone prescription drug plans (PDPs). We observe 1,471 distinct PDPs between 2013 and 2017. 441 plans are present throughout the period, while 420 enter and 782 exit. Towards the end of the period, plan consolidation due to mergers led the number of plans to decline (-15% between 2016 and 2017), a tendency started around 2015. At the same time, differentiation increased, with several insurers offering at least one basic plan and one enhanced plan per region. Plans differ along several dimensions, mainly premiums and generosity (deductible, cost sharing, and coverage). On average, premiums remained approximately constant or decreased slightly throughout the period.

Formularies covered on average 1,123 drugs in 2013 (around 49% generics) and 1,180 in 2017 (around 53% generics), slightly expanding over time as a consequence of both branded and generic entry. Five-tiered formularies became the norm in later years: they usually include two generic tiers (preferred and non-preferred), two brand tiers (preferred and non-preferred), and a single specialty drug tier. Since 2017, CMS allowed the non-preferred brand tier to be replaced by a less restrictive non-preferred drug tier, which could include also generics. Some plans include a sixth tier for injectable drugs.

2.3 Tier Upgrading for Generics

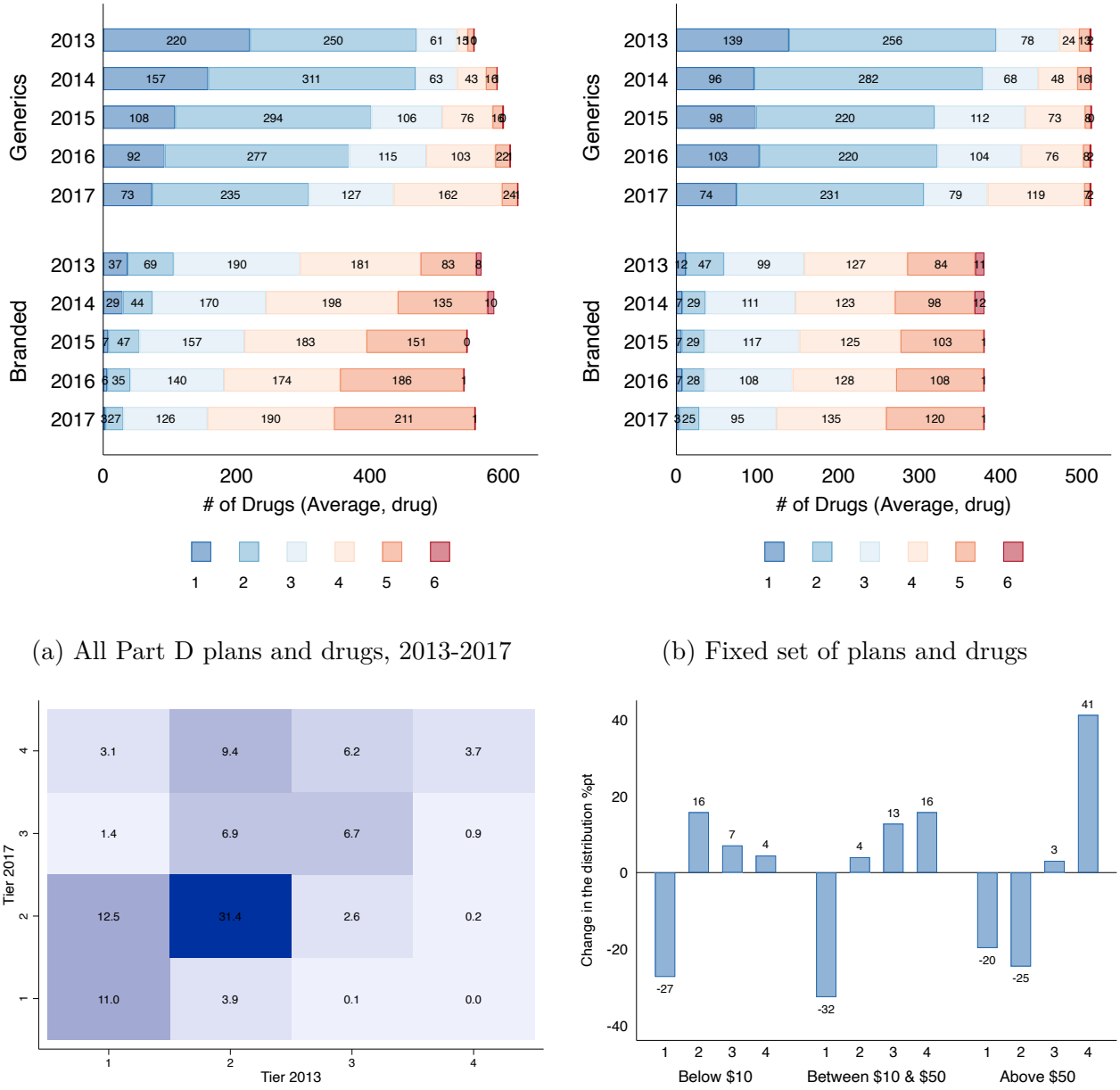
Figure 2.1 documents a systematic shift of generics from lower to higher tiers in the period 2013-2017. Figure 2.1(a) illustrates how, in 2013, around 90% of all generics across plans belonged to tier 1 or 2, with 220 distinct drugs in tier 1. By 2017, only around half of all generics were in tier 1 or 2. Tier 1 had on average shrunk in size to 73 distinct drugs, losing roughly two thirds of the generics. Higher tiers, associated to higher cost sharing, saw an influx of drugs, both branded and generics. Although some of these were new and expensive, which could justify higher cost sharing, most of the shift is attributable to older generics, predominantly upgraded to tier 2 (non-preferred generics) (Figure 2.1(b)). The bottom panels of Figures 2.1(a) and (b) conversely show that tier shifting for branded drugs was limited. The number of branded drugs in tiers 3 and 4 remained stable, while the increase in the size of tier 5 was mostly attributable to the entry of expensive branded drugs. Indeed, in more than 90% of the plans, tier 5 is a specialty tier restricted to drugs with a list price for 30-day supplies above $600. Insurers were allowed to offer a sixth tier for "selected care" drugs associated with a $0 copayment. As illustrated by Figures 2.1(a) and 2.1(b), very few plans used this sixth tier. Overall, this suggests that generic rather than branded drugs tended to be upgraded to higher tiers in 2013-2017, and this will be the focus of the paper.

Figure 2.1(c) shows the transition probability of a drug with a certain tier in 2013 (x-axis) to another tier in 2017 (y-axis). Nearly 40% of the drugs available in both years were moved up by one or two tiers. Conversely, transitions to lower tiers were limited (fewer than 8% of drugs). Interestingly, many generics were also available in tier 3 and 4, mainly devoted to branded drugs, more so in the last years: cheap generics (priced $10-50) increased by 28% their presence in tiers 3 and 4 (non-preferred drugs) (Figure 2.1(d)). This pattern also holds for the most prescribed generic drugs, the top 10 prescribed classes, and those for which CMS mandates that all drugs should be covered in all formularies. This is consistent with Oster and Fendrick (2014), who show that even guideline-recommended medications were often moved out of tier 1 to higher tiers.

2.4 Changes in Out-Of-Pocket Costs

Higher tiers imply higher cost sharing within a plan. In our data, generics moved to higher tiers, on average, experience a doubling in OOP costs: between tier 1 and 2, copayment increased

Figure 2.1: Generics moved to higher tiers



(a) All Part D plans and drugs, 2013-2017

(b) Fixed set of plans and drugs

(c) Tier transition matrix

(d) By price group (based on 2013 list price)

Notes: Figure 1(a) counts the average number of distinct drugs per tier from 2013 to 2017 across plans. Figure 1(b) does the same, fixing the set of plans and drugs as observed in 2013. Average tier size is displayed for generics and branded drugs separately. Figures 1(c) and 1(d) focus on tiers up to the fourth: we exclude drugs with list prices per 30-day supply above \$600 (\$670 in 2017), since they are eligible for a specialty tier intended for expensive drugs. Figure 1(c) maps generic drugs from their tier in 2013 to their tier in 2017. Figure 1(d) shows the change in the distribution of generics across tiers within a certain price group, where price groups are computed on the basis of 2013 list prices.

from around \$1 to more than \$3 in 2017. Generics that are not moved to higher tiers instead display a stable or even decreasing trend in cost sharing since, conditional on tier, plans slightly reduced coinsurance rates. While changes in OOP costs of generics over time are driven by various concurring factors, in this section we attempt to isolate the part due to tier upgrading.

We consider OOP costs for 30-day supply in a preferred pharmacy network in the initial coverage phase of the plan. We denote by ΔOOP_{jpr} and by $\Delta Tier_{jpr}$ the changes, respectively, in out-of-pocket costs and in tier between 2013 and 2017 for drug j in plan p offered in region r, and estimate the following regression:

$$\Delta OOP_{jpr} = \beta_0 + \beta_{1k}\Delta Tier_{jpr} + \delta_j + \delta_p + \varepsilon_{jpr}. \tag{2.1}$$

We allow the coefficient on $\Delta Tier_{jpr}$ to differ across two categories of drugs $k \in \{\text{High Value, Others}\}$. We follow Chandra et al. (2021) and define a drug to be high value if clinical trials demonstrate large mortality benefits.[1] To account for drug-specific changes in OOP costs across plans and regions, we include drug fixed effects δ_j and weigh each observation by the quantity of the corresponding drug-region in 2013. We also include plan fixed effects δ_p, which control for changes in plan design that may have affected *all* drugs in a plan, such as changes in the cost sharing rules for all tiers.

Because formulary design is ultimately a choice of the insurer, we consider $\Delta Tier_{jpr}$ as potentially endogenous and estimate regression (2.1) using an instrumental variable that leverages the relative generosity of plan p for drug j in 2013 compared to the Medicare standard plan. All Medicare Part D plans are required to be at least actuarially equivalent to the government-designed standard plan, which offers a uniform 25% coinsurance rate in the cost-sharing arm above the deductible and below the donut hole. To meet actuarial equivalence, insurers face an incentive to shift to higher tiers those drugs for which the plan is more generous than the standard one, i.e. with a coinsurance *lower* than 25%. For each drug j in plan p, we then compute the difference between the 2013 drug-plan specific coinsurance (as a percentage of the list price) and 25%, and use this as an instrument for $\Delta Tier_{jpr}$.

Table 2.1, column 1, reports the weighted least squares (WLS) estimates of regression (2.1), while column 2 reports the corresponding weighted two-stage least squares (W2SLS) estimates. Compared to generics that did not change tier between 2013 and 2017, those that were moved

[1] High value drugs include statins, ACE inhibitors, β-blockers, thiazide diuretics, calcium channel blockers, angiotensin receptor blockers, diabetes drugs, and inhalants.

Table 2.1: Effect of tier shifting on out-of-pocket costs and list prices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		ΔOOP		Plan level ΔList price			Aggregate list price
	WLS	W2SLS	W2SLS	WLS	W2SLS	W2SLS	WLS
	A. Effect of tier shifting						
$\Delta Tier$	5.929	10.60		3.112	-6.548		
	(0.0273)	(0.109)		(0.0787)	(0.304)		
$1_{\text{High Value}} \times \Delta Tier$			12.62			-3.170	
			(0.179)			(0.507)	
$(1 - 1_{\text{High Value}}) \times \Delta Tier$			9.446			-8.478	
			(0.107)			(0.305)	
$1_{\text{High Value}} \times Year$							-1.172
							(0.0480)
$(1 - 1_{\text{High Value}}) \times Year$							-1.221
							(0.0386)
	B. First Stage						
Dist 25		0.732			0.732		
		(0.00594)			(0.00594)		
$1_{\text{High Value}} \times$ *Dist* 25			0.552			0.552	
			(0.00752)			(0.00752)	
$(1 - 1_{\text{High Value}}) \times$ *Dist* 25			0.858			0.858	
			(0.00674)			(0.00674)	
Weak IV Fstat		15,205	8,197		15,205	8,197	
Observations	195,534	195,534	195,534	195,534	195,534	195,534	102,296
R-squared	0.476	0.074	0.103	0.431	-0.069	-0.068	0.924
Mean Y	6	6		-2.3	-2.3		122.2
Mean Y High value			1.8			.70	53.4
Mean Y (1-High value)			6.8			-2.9	134.4
Mean Δ tier	.8	.8		.8	.8		
Mean Δ tier High value			.4			.4	
Mean Δ tier Others			.8			.8	
Drug FE	✓	✓	✓	✓	✓	✓	✓
Plan FE	✓	✓	✓	✓	✓	✓	
Market FE							✓

Notes: *** p<0.01, ** p<0.05, * p<0.1. Columns 1 to 3 report WLS (weighted least squares) and W2SLS (weighted two-stage least squares) estimates of regression (2.1), where the the first stage regressions are reported in panel B. The sample excludes drugs with list prices per 30-day supply above $600 ($670 in 2017), given their eligibility for a specialty tier intended for expensive drugs. All regressions include drug fixed effects and weigh each observation by the quantity of the corresponding drug-region combination in 2013. "High Value" generics are defined following the classification from Chandra et al. (2021) (see footnote 1). As an instrument for $\Delta Tier_{jpr}$, we leverage the distance between the cost sharing in 2013 and the standard benefit cost sharing rule (25%). Columns 4 to 6 perform similar regressions using ΔList price as dependent variable. Column 7 reports estimates of the list price variation across years using market level data for each generics from the Medicare Part D Event file.

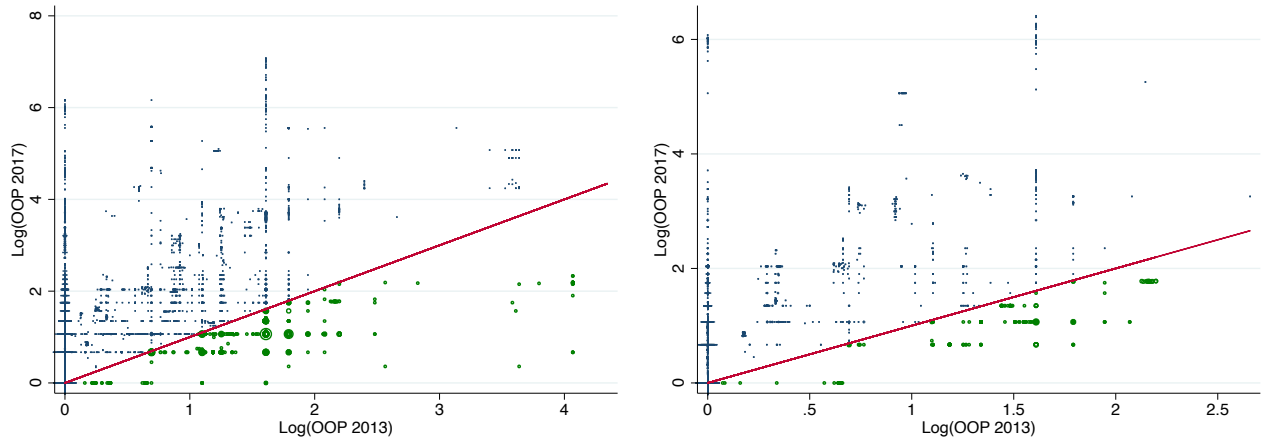
up experienced an increase in OOP costs of \$10.6. The strong first stage results (panel B) and the difference in the estimates between columns 1 and 2 (panel A) suggest that, indeed, formulary design may be endogenous and that insurers did not move generics to higher tiers at random, but rather moved those with OOP costs expected to increase less or even decrease over the period (WLS appears to be downward biased). The relative increase in OOP costs associated to generics that were moved to a higher tier is substantial (column 3): \$12.6 for high value generics and \$9.5 for the others, amounting to a 6 times increase for high value generics and a 0.4 times increase for the other generics.[2] In columns 4 (WLS), 5 and 6 (W2SLS) of Table 2.1, we estimate a regression analogous to (2.1) using changes in the list prices of generics between 2013 and 2017 as the dependent variable. These estimates illustrate that the generics moved to higher tiers were *not* those with relatively increasing list prices. More broadly, column 7 of Table 2.1 shows that the list prices of generics did not increase over the period 2013-2017. These findings suggest that the generalized tier upgrading observed in the period 2013-2017 was not motivated by higher costs of generics for insurers.

Allowing for drug and plan choice adjustments, OOP still increased. In response to increasing OOP costs for those generics that were moved to higher tiers, enrollees may have re-optimized their drug or plan choices. First, some may have switched to therapeutic substitutes at the ATC4 level that were more generously covered. Yet, while in 2013 40% of drug classes (ATC4) had an alternative in tier 1, only 17% had one in 2017. In addition, some enrollees may have switched to plans with lower cost sharing for the drugs that treat the same condition. Although the existing literature has shown that this behavior is not common and plan choices display significant inertia, we allow for this possibility and compute the lowest OOP cost across the plans available in a given Part D market in 2013 and 2017 for a hypothetical enrollee who purchased a drug with a generic available in 2013.

Figure 2.2 summarizes these results. Despite high dispersion, Figure 2.2 clearly shows an upward trend in the OOP costs needed to purchase a 30-day supply of a generic drug. In Figure 2.2(a), more than 90% of drug-market combinations are on or above the 45 degree line (of which 70% are strictly above). Allowing for broader substitution at the class level, in Figure 2.2(b) more than 95% of the combinations are on or above the 45 degree line (of which 50% are strictly above). Hence, tier upgrading increased OOP costs even for the most savvy enrollees. In section 2.5, we enrich this exercise on the basis of the estimated effect of tier upgrading on

[2]From column 3, Table 2.1, we compute 6 and 0.4 as, respectively, $\frac{12.62-1.8}{1.8}$ and $\frac{9.45-6.8}{6.8}$.

Figure 2.2: Minimum out-of-pocket costs increase



(a) Generic name level

(b) ATC4 level

Notes: The figure plots the minimum OOP cost in each Medicare Part D market in 2017 against its corresponding value in 2013. The minimum is computed in panel (a) at the molecule level and in panel (b) at the ATC4 level, allowing for substitution to the cheapest molecule within a drug class. We focus on molecules for which at least one generic was available in 2013.

purchased quantities and OOP spendings.

2.5 Changes in Quantity and Spending

To shed light on the consequences of tier upgrading of generics on enrollees, we first investigate its effect on the demand for generics.

We define quantity Q_{jrt} as the 30-day supply of a generic drug j, Medicare Part D region r, and year t. We estimate the following regression:

$$\log(Q_{jrt}) = \alpha_0 + \alpha_{1k} Tier_{jrt} + X_{jrt}\alpha_2 + \zeta_j + \zeta_r + \zeta_t + \nu_{jrt}, \tag{2.2}$$

where $k \in \{\text{High Value, Others}\}$. We compute $Tier_{jrt}$ as the weighted average of the tiers which generic drug j belongs to in each plan, using enrolment weights, and round it to the nearest integer.[3] X_{jrt} includes as controls generic drug coverage in the market (share of plans), entry at the ATC4 level, coverage of the branded alternative (share of plans), and size of the market (in number of enrollees).

Similar to the price regression (2.1), $Tier_{jrt}$ in regression (2.2) is likely to be endogenous: generic drugs with larger unobserved components of demand may be placed in higher tiers to

[3]Using non-rounded measures of average tier at the market level yields similar estimates.

monetize a positive demand shock. We build a Hausman-type instrument for $Tier_{jrt}$ (Hausman (1996); Nevo (2001)) as the average tier of generic j in regions other than r. This is expected to correlate with $Tier_{jrt}$ due to the common components of cost incurred by insurers for offering generic j in their plans independently of the region (e.g., bargaining of insurers with drug manufacturers), but should be uncorrelated with the local demand shock ν_{jrt}. Controlling for drug, region, and year fixed effects, the validity of this instrument rests on the absence of residual aggregate demand shocks correlated both with ν_{jrt} and with $\nu_{jr't}$, $r' \neq r$, and that led to a national-wide tier shifting for generic j across plans.

Given that insurers may use the same formulary in several Part D regions, generic j's tier in region r' could correlate with ν_{jrt} through a common formulary used both in regions r' and r. Hence, we exclude common formularies from the instrument. Moreover, we check the robustness of our results by adding region-specific time trends and drug class (ATC3)-specific time trends meant to capture residual nation-wide demand shocks that may invalidate our instrument.

Table 2.2: The Effect of Tier on Demand, 2013-2017

	(1) OLS	(2) 2SLS	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
	A. Effect of tier on $\log(Q_{jrt})$					
$Tier$	-0.0324	-0.173		-0.204	-0.155	-0.191
	(0.00466)	(0.0249)		(0.0299)	(0.0318)	(0.0384)
$1_{\text{High Value}} \times Tier$			-0.318			
			(0.0732)			
$(1 - 1_{\text{High Value}}) \times Tier$			-0.165			
			(0.0237)			
	B. First Stage					
Average of j's tier in $r' \neq r$		0.230	0.230	0.209	0.179	0.166
		(0.00512)	(0.00512)	(0.00500)	(0.00484)	(0.00484)
Weak IV Fstat		2,012	174.8	1,752	1,370	1,176
Observations	83,232	83,157	83,157	83,157	83,157	83,157
R-squared	0.968	0.799	0.798	0.726	0.806	0.733
Drug FE	✓	✓	✓	✓	✓	✓
Region FE	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓			
Region FE $\times$ $Year_t$				✓		✓
ATC3 FE $\times$ $Year_t$					✓	✓

Notes: *** p<0.01, ** p<0.05, * p<0.1. The table reports OLS and 2SLS estimates of various versions of regression (2.2), where the the first stage regressions are reported in panel B. Column 1 reports OLS estimates. Columns 2 to 6 instead report 2SLS estimates using the average tier of generic j in regions other than r as an instrument for $Tier_{jrt}$. Column 3 reports separate estimates for high value and other generics, with "high value" generics defined according to the classification from Chandra et al. (2021) (see footnote 1). Columns 4 to 6 further include region-specific and drug class-specific time trends to control for national-wide aggregate demand shocks that would invalidate our instrument.

Column 1 of Table 2.2 reports the OLS estimates of regression (2.2), while columns 2-6 the corresponding 2SLS estimates. A comparison of the OLS estimates in column 1 with the 2SLS estimates in column 2 (around 5 times more negative, with a strong first stage in panel B) confirms that $Tier_{jrt}$ may be the outcome of endogenous insurers' decisions and is unlikely to be independent of unobserved demand shocks. Reassuringly, the estimated effect of $Tier_{jrt}$ on $\log(Q_{jrt})$ appears to be robust to the inclusion of region-specific and drug class-specific time trends (columns 4-6). In line with Chandra et al. (2021), column 3 illustrates that the effect of $Tier_{jrt}$ on $\log(Q_{jrt})$ is estimated to be stronger for high value (-0.318) than for other generics (-0.165).

To put these estimates into perspective, we combine them with those from Table 2.1 and compute the elasticity of demand with respect to the increase in OOP costs implied by tier

upgrading, $\epsilon = \frac{\Delta Q}{Q}\frac{OOP}{\Delta OOP} = (\exp(\alpha_{1k}) - 1)\frac{OOP}{\beta_{1k}}$. The average estimated elasticity among all generics is $\epsilon_{\text{all}} = -0.18$, which is in line with the estimate of -0.14 by Einav et al. (2018) for the Medicare Part D population close to the coverage gap. By considering high value and other generics separately, we obtain $\epsilon_{\text{high value}} = -0.12$ and $\epsilon_{\text{others}} = -0.21$, which suggests that the demand for high value generics is *less* elastic than that for the other generics. This is consistent with the lower elasticities reported by Einav et al. (2018) for chronic and maintenance drugs, a category that includes high value drugs.

Counterfactuals. The above evidence suggests that enrollees paid more for generics due to tier upgrading and that higher cost sharing caused a decrease in their purchased quantities. In this section, building on the idea of Figure 2.2 and on our regressions, we evaluate these changes from the perspective of enrollees' total spending.

To approximate the total annual OOP spending (OOPS) per-enrollee for purchases of generic drugs, we use aggregate Medicare Part D prescriber quantities and cost sharing rules at the plan level, weighted by plan enrolment:

$$\begin{aligned} OOPS_{rt} &= \frac{1}{N_{rt}} \sum_{j \in \mathcal{J}_{rt}} \left(Q_{jrt} \times \sum_{p \in \mathcal{P}_{jrt}} m_{jprt} \times OOP_{jprt} \right) \\ &= \frac{1}{N_{rt}} \sum_{j \in \mathcal{J}_{rt}} \left(Q_{jrt} \times OOP_{jrt} \right), \end{aligned} \tag{2.3}$$

where, for each region r and year t: N_{rt} is the total number of Medicare Part D enrollees; $\mathcal{J}_{rt}$ is the set of generic drugs available; $\mathcal{P}_{jrt}$ is the set of plans that include generic j; m_{jprt} is the observed market share of plan p (among those covering j); Q_{jrt} is the purchased quantity of generic j; OOP_{jprt} is the OOP cost of generic j in plan p for a 30-day supply in a preferred pharmacy during the initial coverage period; and OOP_{jrt} is the weighted average of OOP_{jprt} across plans covering j.

We ask the following question: how did tier upgrading between 2013 and 2017 affect OOP spending in 2017? Relying on the estimates of regressions (2.1) and (2.2) from Tables 2.1 and 2.2 (column 3 for both), we simulate OOPS (2.3) in 2017 given the tier structure observed in 2013. We exclude from the analysis the generics observed to enter or exit in 2017 as well as those for which the tier did not change between 2013 and 2017, as for these drugs we would not be able to compute any counterfactual change. The generics included in our sample represent 74% of all generics and 68% of total purchased quantities of generics in 2017. To simplify the

simulations, we assume that the plan market shares in (2.3) remained constant at their observed levels in 2017 when altering the tier structure to that observed in 2013.[4]

To isolate the contribution of tier shifting to the change in cost sharing between 2013 and 2017, we use the estimates from regression (2.1) (column 3) and compute counterfactual OOP costs in 2017 as:

$$OOP^{c}_{jpr17} = \widehat{\Delta OOP}_{jpr} - \widehat{\beta}_{1k}\Delta Tier^{c}_{jpr} + OOP_{jpr13}, \tag{2.4}$$

where $\widehat{\Delta OOP}_{jpr}$ is the fitted value of regression (2.1), $\widehat{\beta}_{1k}$ is the corresponding estimate of β_{1k}, $\Delta Tier^{c}_{jpr}$ is a counterfactual shift in tier (e.g., 0 in the absence of a change or $-n$ for a decrease of n tiers) and OOP_{jpr13} is the OOP cost observed in 2013. We denote by OOP^{c}_{jr17} the weighted average of OOP^{c}_{jpr17} across plans using plan enrolment weights, m_{jpr17}.

To account for the effect of tier shifting on demand for generics, we use the estimates of regression (2.2) (column 3) and compute counterfactual purchased quantities in 2017 as:

$$Q^{c}_{jr17} = \widehat{Q}_{jr17} \exp\left(\widehat{\alpha}_{1k} \times \Delta Tier^{c}_{jpr}\right), \tag{2.5}$$

where $\widehat{Q}_{jr17}$ is the fitted value for 2017 of regression (2.2), $\widehat{\alpha}_{1k}$ is an estimate of α_{1k}, and $\Delta Tier^{c}_{jpr}$ is a counterfactual shift in tier.

Combining (2.3), (2.4), and (2.5), we compute the total counterfactual change in OOP spending in 2017 as if tier shifting between 2013 and 2017 had not taken place:

$$\Delta OOPS^{c}_{r17} = \frac{1}{N_{r17}} \sum_{j \in \mathcal{J}_{r13}} \left(\widehat{Q}_{jr17}\widehat{OOP}_{jr17} - Q^{c}_{jr17}OOP^{c}_{jr17}\right),$$

where $\mathcal{J}_{r13}$ denotes the set of generics available both in 2013 and 2017 and that changed tier and $\widehat{OOP}_{jr17}$ is computed on the basis of (2.4) evaluated at $\Delta Tier^{c}_{jpr} = 0$ as $\widehat{OOP}_{jr17} = \sum_{p \in \mathcal{P}_r} m_{jpr17} \times \left(\widehat{\Delta OOP}_{jpr} + OOP_{jpr13}\right)$. In words, $\Delta OOPS^{c}_{r17}$ is the difference between $\widehat{Q}_{jr17}\widehat{OOP}_{jr17}$, the predicted OOPS in 2017 given the tier structure in 2017, and $Q^{c}_{jr17}OOP^{c}_{jr17}$, the counterfactual OOPS in 2017 given the tier structure in 2013. In evaluating $\Delta OOPS^{c}_{r17}$, we use the model's prediction of the factual OOPS ($\widehat{Q}_{jr17}\widehat{OOP}_{jr17}$) rather than the corresponding observed value ($Q_{jr17}OOP_{jr17}$) to obtain a cleaner comparison with the counterfactual OOPS ($Q^{c}_{jr17}OOP^{c}_{jr17}$), which can only be predicted by the model.

[4] In general, this assumption could be relaxed and the variable m_{jprt} in equation (2.3) endogenized by estimating a plan level choice model along the lines of Decarolis et al. (2020).

While $\Delta OOPS^{c}_{r17}$ measures the total counterfactual change in OOPS due to higher OOP costs and implied demand responses, we disentangle the contribution of these channels by decomposing $\Delta OOPS^{c}_{r17}$ in two parts. We compute the part of the counterfactual change in OOPS due to the higher OOP costs implied by tier upgrading between 2013 and 2017 as:

$$\Delta OOPS^{oop}_{r17} = \frac{1}{N_{r17}} \sum_{j \in \mathcal{J}_{r13}} Q^{c}_{jr17} \left(\widehat{OOP}_{jr17} - OOP^{c}_{jr17} \right).$$

We then compute the part of $\Delta OOPS^{c}_{r17}$ due to the lower purchased quantities implied by tier upgrading between 2013 and 2017 as:

$$\Delta OOPS^{q}_{r17} = \frac{1}{N_{r17}} \sum_{j \in \mathcal{J}_{r13}} \left(\widehat{Q}_{jr17} - Q^{c}_{jr17} \right) \widehat{OOP}_{jr17},$$

such that $\Delta OOPS^{c}_{r17} = \Delta OOPS^{oop}_{r17} + \Delta OOPS^{q}_{r17}$.

Table 2.3 reports the results of the counterfactual, for all drugs (column 1) and separately for drugs in the high value and in the others category. For each category, we further separate the effects for drugs that were moved to higher tiers (column "Up") from the handful that were moved down (column "Down"). Panel A displays the breakdown of drugs in terms of number of distinct products and sales.

Panel B presents the effect of tier shifting between 2013 and 2017 on OOP spending in 2017. Compared to a scenario in which the tier structure had remained the same as in 2013, OOP spending per-enrollee in 2017 would have increased by \$76, accounting for two thirds of the total increase in spending on generics between 2013 and 2017, \$116. This corresponds to an additional \$1.5 billions in OOP spending for the Medicare Part D population in 2017. As expected by their pervasiveness on formularies and their higher prices and tiers, drugs in the "Others" category that were moved to higher tiers are the largest contributors to these estimated increases, although upgraded high value drugs also account for an important share of the change in OOP spending. The total effect on OOP spending is mitigated by a few drugs that were moved to lower tiers between 2013 and 2017 (the negative change in spending in the columns denoted by "Down").

The increase of \$76 can be decomposed in two parts. First, if demand had not decreased in response to the higher OOP costs, OOP spending per-enrollee would have increased by \$100, a third more. Second, the reduction in demand induced by the higher OOP costs would have led to a decrease in OOP spending per-enrollee by \$24, had tier placement not changed compared

to 2013. Thus, despite the low elasticity estimated by our model and the literature, the large increase in cost sharing induced a reduction in consumption, which translated into lower spending as measured by $\Delta OOPS^{q}_{r17}$. The decomposition of the total counterfactual change highlights a trade-off between adherence to therapy and spending and suggests interesting heterogenous effects across drugs and enrollees. An hypothetical enrollee on anti-diabetics, a class with an estimated perfectly inelastic demand to price Einav et al. (2018), would have spent $100 more in 2017 in order to perfectly adhere to therapy. Conversely, an enrollee on anti-hypertensive or β-blockers, two classes with a more elastic demand despite high mortality benefits, would have cut on drugs to reduce spending, increasing their health hazards.

Table 2.3: Counterfactual Change in OOP spending per enrollee

	All	High Value		Others	
		Up	Down	Up	Down
	A. Tier shifting between 2013 and 2017				
Drugs (%)	100	8.9	1.5	60.9	2.6
Quantity (%)	100	13.3	7.6	40.2	6.7
	B. Change in OOP Spending in 2017 ($)				
Total change: $\Delta OOPS^{c}_{r17}$	75.7	15.8	-10.1	75.0	-4.4
Part due to ΔOOP costs: $\Delta OOPS^{oop}_{r17}$	99.5	20.4	-10.5	94.8	-4.7
Part due to ΔQuantities: $\Delta OOPS^{q}_{r17}$	-23.9	-4.7	0.5	-19.8	0.3
Enrolment 2017	19,694,854				

Notes: The table reports the counterfactual per-enrollee change in OOP spending on generics in 2017 due to tier shifting between 2013 and 2017. We restrict the analysis to common drug classes (with more than 100,000 claims over the period) and to generics available both in 2013 and 2017 that changed tier between 2013 and 2017. The counterfactual computations are based on the estimates of regressions (2.1) and (2.2) from Tables 2.1 and 2.2 (column 3 for both), and equations (2.3), (2.4), and (2.5). Panel A shows the distribution of drugs by category (high value and other generics) and tier shifting status between 2013 and 2017. The difference between 100 and the sum of the last four columns refers to drugs that remained in the same tier (not reported in the interest of space). Panel B shows the corresponding counterfactual per-enrollee change in OOP spending on generics in 2017 ($\Delta OOPS^{c}_{r17}$), and its decomposition into the part due to changes in OOP costs ($\Delta OOPS^{oop}_{r17}$) and the part due to changes in quantities ($\Delta OOPS^{q}_{r17}$).

2.6 Welfare Implications

In this section, we discuss some of the welfare implications of the observed tier upgrading of generics from 2013 to 2017. First, we use the adjusted mortality data from Chandra et al. (2021) to perform a back-of-the-envelope calculation of the consequences of cutting on high

value drugs due to tier upgrading on mortality.[5] Our estimates imply that tier shifting caused a 10% reduction in the consumption of these drugs compared to a scenario without tier upgrading, which would translate into mortality effects of 3.7% for ACE inhibitors, thiazide diuretics, and angiotensin-receptor blockers and of 8.2% for β-blockers. To scale these to the relevant populations, we use the shares of eligible enrollees by class from Chandra et al. (2021).[6] Summing up across the four high values classes, accounting for the different levels of utilization in the Part D population across them, we obtain an overall increase in mortality of 5.4%.

In addition to significant increases in spending and possible adverse health consequences due to reduced adherence to therapy, tier upgrading may have also affected welfare by distorting the optimal cost sharing required to balance risk protection and moral hazard. In line with the theory of optimal insurance (Feldstein (1973); Besley (1988)), Einav et al. (2018) document that, in the period 2007-2011, drugs with higher demand elasticity to price (that is, more subject to moral hazard) had higher cost sharing due to placement in higher tiers. Yet, their results suggest lower than optimal average coverage. In this sense, tier upgrading in the period 2013-2017 may have narrowed the gap and brought further reductions in moral hazard. We investigate this possibility and ask whether drugs with higher demand elasticity were more likely to be moved to higher tiers. By matching our data to the estimated demand elasticities for common drugs by Einav et al. (2018), we find that the generics moved to higher tiers between 2013 and 2017 tended to have *lower* rather than higher estimated demand elasticity. This suggests that tier upgrading, by moving the cost sharing structure further away from optimality, may have negatively affected welfare beyond reductions in consumption and increases in spending.

2.7 Conclusions

In this paper, we document that, in the period 2013-2017, Medicare Part D plans systematically upgraded many generic drugs to higher tiers, with an implied average increase in OOP spending per-enrollee of $76, or $1.5 billion for the entire Medicare Part D population.

[5] Among the high value drugs, we exclude antidiabetics, calcium channel blockers, and inhalants, as their demand is estimated to be almost perfectly inelastic by Einav et al. (2018). We are possibly underestimating the quantity effect on the remaining high value drugs, as our demand model has a single coefficient for all high value drugs. We also exclude statins, since one of the main drugs in this class, atorvastatin, was moved to a lower tier.

[6] We do not use the fraction of beneficiaries by class in 2013 from our data, since it would significantly overestimate the relevant population: our aggregate data only report the number of beneficiaries at the drug level, so we would count more than once those that use more than one drug in the same class in the same year. The shares reported in Chandra et al. (2021) include also branded drugs, so may be overestimated, but are lower than those in our data and provide a reasonable approximation.

The formulary is a key strategic variable for insurers and changes in tier design are consistent with the nature of competition in Medicare Part D. As documented by an extensive literature, premiums are the most salient feature on which consumers compare plans and insurers compete fiercely over them. Moreover, increased regulatory scrutiny has contributed to a remarkable stability of premiums over the past years, despite significant growth in spending for Part D's catastrophic benefits. Formularies, on the other hand, remain relatively unregulated, inducing insurers to exploit this flexibility to increase profits.

While sizeable, our estimates are likely to be a lower bound for the total increase in both private and public spending. A lower effective OOP difference between brand-name and generics may shift some consumption towards the former (Dafny et al. (2017); Dubois and Lasio (2018)) and increase OOP even further. In turn, higher OOP costs may push some enrollees to the catastrophic threshold at a faster pace, increasing total expenditures for Medicare. Additionally, lower utilization of high value drugs may lead to a surge in hospitalization (Chandra et al. (2010, 2021)), with higher costs for both the enrollees and Medicare.

Our results hinge on a number of assumptions. First, given the aggregate nature of our data, we do not consider substitution across drugs, which may affect health outcomes and spending. Also, our counterfactuals consider plan market shares to be unaffected by tier shifting and thus cannot capture potential heterogeneous effects across enrollees. Finally, we provide back-of-the-envelope welfare computations that rely on estimates from the literature. All these limitations could be addressed by developing a richer model estimated with individual-level data, which we leave for future research.

Bibliography

Abaluck, J. and Gruber, J. (2011). Choice inconsistencies among the elderly: Evidence from plan choice in the medicare part d program. *American Economic Review*, 101(4):1180–1210.

Abaluck, J. and Gruber, J. (2016). Evolving choice inconsistencies in choice of prescription drug insurance. *American Economic Review*, 106(8):2145–84.

Armstrong, M. and Chen, Y. (2009). Inattentive Consumers and Product Quality. *Journal of the European Economic Association*, 7(2-3):411–422.

Avalere (2018). Generic drugs in medicare part d. trends in tier structure and placement. Technical report.

Berndt, E., Conti, R., and Murphy, S. (2017). The landscape of US generic prescription drug markets, 2004-2016. NBER Working paper No. 23640.

Besley, T. J. (1988). Optimal reimbursement health insurance and the theory of ramsey taxation. *Journal of Health Economics*, 7(4):321–336.

Carlin, B. I. (2009). Strategic price complexity in retail financial markets. *Journal of Financial Economics*, 91(3):278–287.

Chandra, A., Flack, E., and Obermeyer, Z. (2021). The health costs of cost-sharing. Technical report.

Chandra, A., Gruber, J., and McKnight, R. (2010). Patient cost-sharing and hospitalization offsets in the elderly. *American Economic Review*, 100(1):193–213.

Dafny, L., Ody, C., and Schmitt, M. (2017). When discounts raise costs: The effect of copay coupons on generic utilization. *American Economic Journal: Economic Policy*, 9(2):91–123.

Dalton, C. M., Gowrisankaran, G., and Town, R. J. (2019). Salience, Myopia, and Complex Dynamic Incentives: Evidence from Medicare Part D. *The Review of Economic Studies*, 87(2):822–869.

Decarolis, F. (2015). Medicare part d: Are insurers gaming the low income subsidy design? *American Economic Review*, 105(4):1547–80.

Decarolis, F., Polyakova, M., and Ryan, S. P. (2020). Subsidy design in privately provided social insurance: Lessons from medicare part d. *Journal of Political Economy*, 128(5):1712–1752.

Dubois, P. and Lasio, L. (2018). Identifying industry margins with price constraints: Structural estimation on pharmaceuticals. *American Economic Review*, 108:3685–3724.

Duggan, M. and Scott Morton, F. (2010). The effect of medicare part d on pharmaceutical prices and utilization. *American Economic Review*, 100(1):590–607.

Dusetzina, S. B., Jazowski, S., Cole, A., and Nguyen, J. (2019). Sending the wrong price signal: Why do some brand-name drugs cost medicare beneficiaries less than generics? *Health Affairs*, 38(7):1188–1194. PMID: 31260351.

Einav, L., Finkelstein, A., and Polyakova, M. (2018). Private provision of social insurance: drug-specific price elasticities and cost sharing in medicare part d. *American Economic Journal: Economic Policy*, 10(3):122–153.

Ellison, G. and Ellison, S. F. (2009). Search, obfuscation, and price elasticities on the internet. *Econometrica*, 77(2):427–452.

Feldstein, M. S. (1973). The welfare loss of excess health insurance. *Journal of Political Economy*, 81(2, Part 1):251–280.

Fix, A., J. B. S. D. and Whorley, M. (2022). 57% of generic drugs are not on 2022 part d generic tiers. Technical report, Avalere Health.

Fleitas, S. (2020). Who benefits when inertia is reduced? competition, quality and returns to skill in health care markets. Technical report.

Handel, B. R. and Kolstad, J. T. (2015). Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8):2449–2500.

Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In *The economics of new goods*, pages 207–248. University of Chicago Press.

Ho, K., Hogan, J., and Scott Morton, F. (2017). The impact of consumer inattention on insurer pricing in the medicare part d program. *The RAND Journal of Economics*, 48(4):877–905.

Lavetti, K. and Simon, K. (2018). Strategic formulary design in medicare part d plans. *American Economic Journal: Economic Policy*, 10(3):154–92.

Marzilli Ericson, K. M. (2014). Consumer inertia and firm pricing in the medicare part d prescription drug insurance exchange. *American Economic Journal: Economic Policy*, 6(1):38–64.

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

Olssen, A. L. and Demirer, M. (2021). Drug rebates and formulary design: Evidence from statins in medicare part d. In *Massachusetts Institute of Technology Working Paper*.

Oster, G. and Fendrick, A. (2014). Is all "skin in the game" fair game? the problem with "non-preferred" generics. *The American journal of managed care*, 20:693–5.

Starc, A. and Town, R. J. (2020). Externalities and benefit design in health insurance. *The Review of Economic Studies*, 87(6):2827–2858.

Chapter 3

Biologic Drugs and Learning-by-Doing

Biologic drugs – complex treatments derived from living cells – now represent nearly half of pharmaceutical sales in developed countries. Unlike traditional molecules, biologics exhibit substantial marginal costs and the complexity of the manufacturing processes leaves scope for significant efficiency gains through learning-by-doing. This paper quantifies learning-by-doing in biologic drug manufacturing, where marginal costs decline as cumulative production increases. Using product-level sales data across 25 countries, we document a negative correlation between past sales and current prices. Focusing on TNF-α inhibitors, we estimate a structural model of demand, pricing, and production. We find that biosimilar marginal costs declined by 46% on average since market entry. Learning spillovers arise from concentrated production: delays in biosimilar entry in Canada raised EU marginal costs by 2.4% 15 months after market entry. Counterfactuals show that the timing and geography of biosimilar uptake matter for global cost efficiencies.

Keywords: Pharmaceuticals, Biologic Drugs, Learning-by-doing.

JEL codes: L20, I10

3.1 Introduction

Pharmaceutical markets have traditionally featured high Research & Development (R&D) costs followed by low marginal costs of producing successful products. Over time, the industry shifted to more complex treatments, made from living cells, denominated biologic drugs. These elaborated therapies are increasingly prevalent in developed countries, representing 40 and 45% of pharmaceutical sales in the European Union and the US in 2023.[1] They are also more expensive than traditional molecules. While patent protection rewards innovation by granting temporary monopoly power, the accessibility and sustainability of healthcare systems rely on the competition exerted by the entry of copycat products following patent expiration. In the case of biologic drugs, this process is altered by their living-cell origin, along with the complex manufacturing process. The latter makes production experience valuable, leaving scope for returns to experience to decrease future marginal costs and prices. This is particularly important for healthcare systems, as biologics can have substantial marginal costs. Moreover, traditional drug pricing schemes may not be optimal when marginal costs evolve dynamically – making it crucial to measure potential efficiency gains for informed public policy.

To study these questions, we assemble publicly available data on the production process of biologic drugs and combine this information with proprietary country-level data on drug consumption. Relying on drug-level pharmaceutical sales data for 25 countries, we document the negative correlation between past sales and current prices across thirteen biologic molecules. Then, we quantify learning-by-doing in the production of Tumor Necrosis Factor (TNF)-α inhibitors, which are biologic agents to treat rheumatoid arthritis, spondylarthritis, plaque psoriasis, and chronic inflammatory bowel diseases. We rely on a structural model of TNF-α inhibitors demand, pricing, and production. More precisely, we overcome the lack of product-level cost information by estimating a model of drug demand and pricing. The marginal cost estimates are used to recover the primitives of the production model. The estimates suggest that the marginal cost of biosimilars decreases by 10.3% every time experience doubles. Although small, this number needs to be put into perspective in relation to the production unit we consider: one daily dose of a drug. On average, the marginal cost of TNF-α inhibitors biosimilars decreased by 46% since market entry. The concentration of biologic drug production processes induces spillovers across biologic drug consumption in different countries. As a

[1]Source: Assessing the biosimilar void in the U.S. IQVIA (2025) The Impact of Biosimilar Competition in Europe (IQVIA January 2025)

consequence, demand in each country exerts positive intertemporal spillovers on the production efficiency of biologic drugs, reducing marginal costs of production in the future. Thus, delays in the entry of TNF-α inhibitors biosimilars in Canada increased marginal costs in the EU by 2.4% 15 months after entry.

This paper quantifies learning-by-doing in the biologic drugs production process, where learning-by-doing is characterized by a decrease in marginal costs as cumulative past production increases. To that end, we combine four sources of data proving information on biologic drugs' (*i*) characteristics, (*ii*) consumption since market entry, (*iii*) production facilities, and (*iv*) changes in the production process post-market entry. To overcome the lack of information about drug-level production costs over time, we develop a model of biologic drug demand, pricing, and production to estimate returns to experience without production data. Our approach is similar to Khmelnitskaya et al. (2024) who evaluates scale and scope economies in the US beer industry. As shown by De Loecker and Scott (2024) in the case of the US brewing industry, using production or demand data to estimate markups produce similar estimates. As production data for pharmaceuticals and biologics are not available, estimating marginal costs of biologics needs to use a demand and pricing approach. On the supply side, the production of biologic drugs across countries is concentrated in a few factories, subject to learning-by-doing from past productions and economies of scale. In this context, the pricing game is inherently difficult to solve, preventing us from recovering marginal costs from the observed pricing equilibrium.[2] We overcome this challenge by estimating marginal costs based on pricing behavior in a small market: Sweden. The units consumed in the Swedish market are produced in the same facilities as those in other EU markets, and thus are subject to similar marginal costs of production. Yet, the limited market size suggests that volumes sold in Sweden contribute marginally to learning-by-doing, limiting the scope for dynamic considerations when setting drug prices in Sweden. We can thus recover marginal costs using standard Bertrand Nash pricing conditions in static oligopolies. These estimates enable the recovery of the primitives of the production model using an instrumental variable approach to overcome endogeneity concerns.

In preliminary counterfactual analysis, we use the primitives of the production model to compute marginal costs under varying levels of experience. In particular, we compute marginal costs that would have prevailed in the absence of biosimilar consumption in France, Germany, and the United Kingdom. Country-level biosimilar penetration matters for cost efficiencies in

[2]Solving and estimating the primitives of the pricing game is beyond the scope of this version of the paper. In future iterations, we plan to improve the paper in this dimension.

other countries, and the size of these spillovers varies with the size of the own-country market. Without entry in the German market, biosimilar costs would have been 4% higher at the end of the sample period. This effect is heterogeneous throughout the product's lifespan, depending on how fast patients adopted biosimilars in this country versus others. The second counterfactual investigates the consequences of delaying biosimilar entry. These delays are rather common, especially in the US and Canada, driven by delays in regulating biosimilar entry and patent litigations. While learning-by-doing suggests that delaying biosimilar entry results in lower marginal costs upon entry, these delays slow down efficiency gains in other countries. The cost of biosimilars when entering the Canadian market was 64.5% lower, yet marginal costs would have been up to 7.4% lower 15 months after entry, had biosimilars entered in Canada at the same time as they entered the EU.

Related literature The primary contribution of this paper is to develop an empirical framework for studying learning-by-doing, relying on sales data. Primarily, this project builds on the theoretical and empirical literature studying learning-by-doing, summarized by Thompson (2010) and Thompson (2012). In the empirical literature, Benkard (2000) and Benkard (2004) study learning-by-doing in the context of commercial aircraft. Relying on production costs data, the author finds that pricing below static marginal cost can be explained by learning-by-doing and forgetting. As a result, Benkard (2004) finds that a dynamic oligopoly model with dynamic marginal costs rationalizes some features of the commercial aircraft market structure. Using detailed data on car assembling defects and workers, Levitt et al. (2013) documents that the knowledge stock acquired through learning-by-doing is not retained by workers and becomes quickly embodied in the organization of production. Considering learning-by-doing and organizational forgetting in a model of price competition, Besanko et al. (2010) highlights that forgetting can lead to aggressive pricing behavior, affecting industry concentration in the long run. This paper also builds on the theoretical literature analyzing the implications of learning-by-doing (Arrow (1962), Cabral and Riordan (1994), Fudenberg and Tirole (1983)).

Our paper differs from this literature in several aspects. First, unlike Benkard (2004), the primary strategic variable in the context of a pharmaceutical market is price. Unlike quantities, when firms compete in prices, a change in the strategic variable of your competitors affects your quantities, hence your state-to-state transition (Doraszelski and Pakes (2007)). Second, while the existing empirical literature has relied on production data, we combine sales data with an

empirical model of drug pricing to estimate returns to experience. As a result, the methodology used in this paper is similar to that of Khmelnitskaya et al. (2024), which combines a micro-founded model of production with sales data to estimate economies of scale and scope in the beer industry. One important difference in their context is that the type of cost efficiencies they are modelling does not involve dynamic considerations on the firm side, allowing them to recover marginal costs from a static Bertrand-Nash pricing game at each period.

Finally, this paper also contributes to the literature on biosimilars and their impact on market outcomes. Analyzing the impact of biosimilar entry in Europe, Scott Morton et al. (2018) find that price reductions and market penetration depend largely on market size and demand-side institutions. In the US, Feng et al. (2024) finds that branded drug manufacturers are more likely to compete with biosimilars than they do with generics to maintain their volume. They rationalize this behavior by the lower number of entrants, the lack of perfect equivalence, and the perceived difference in quality on the demand side between branded drugs and biosimilars.

The paper proceeds as follows. Section 3.2 provides background information on biologic drug production and describes the data. Section 3.3 provides empirical evidence supporting learning-by-doing in biologic drug production. Section 3.4 presents a model of drug production, pricing, and demand. Section 3.5 estimates the empirical model. Section 3.6 presents the counterfactual scenarios, and Section 3.7 concludes.

3.2 Setting and Data

3.2.1 Biologic drugs and biosimilars

Biologic drugs are large and complex molecules produced in living organisms, which allows them to exhibit a high degree of structural complexity. While the molecular mass of conventional drugs typically lies around a few hundred grams per mole, the molecular mass of biologic drugs can range from 6,000 grams per mole (for insulins, for example) to 150,000 grams per mole (for monoclonal antibodies). The production process for prescription drugs generally consists of two main steps: synthesizing the active ingredient and formulating the drug into its final product form. In the context of biologics, manufacturing the drug substance relies on cell line culture and several purification procedures. The second step involves combining the active ingredient with excipients and packaging the drug. Biologic drugs are sensitive to changes

in the production process, making it challenging to produce a sufficiently uniform product. Manufacturing biologic drugs requires expertise and experience, and manufacturers face a steep learning curve (Blackstone and Joseph (2013)). Yet, regulatory agencies such as the EMA and the FDA mandate that drug manufacturers submit any variations in the production process, thereby tracking changes in the production process post-market approval. These changes can trigger many aspects of the products, from the manufacturing and testing of the active substance to the package leaflet. In the EU, minor variations can be implemented before they are declared to the EMA (the "Do and tell" procedure), while others must be declared ex-ante (the "Tell, wait and do"), and others require prior approval before being implemented. In the first case, the implementation can be fast as it does not require any prior authorization from the agency. "Tell, wait, and do" procedures are usually approved within one to three months, and five to twelve months in the latter case.[3]

Both the production process of biologic products and the nature of the molecule itself affect the extent of competition after patent expiry. First, the complexity of the production process requires specialized facilities. Second, the molecule is produced from living cells, and each stem cell is unique. As a result, unlike small molecules, biologic drugs cannot be perfectly replicated by competing drug manufacturers. Hence, the equivalent of generic drugs for biologic molecules is denominated Biosimilars. This feature affects the entry costs of potential biosimilars, which are estimated to be between $100 million and $250 million, ten times larger than those for generic drugs.[4] On the other hand, originator facilities may be older, such that biosimilar manufacturers relying on newer technology can face lower baseline manufacturing costs (Blackstone and Joseph (2013)).

In this context, learning-by-doing links past production to current cost efficacy, inducing dynamic consideration by drug manufacturers as well as spillovers across geographical markets. Indeed, the experience gained in one market can enhance effectiveness in other markets. The first objective of this paper is to measure the impact of learning-by-doing on the marginal costs of biologic drugs and biosimilars.

[3] The guidelines applicable from 2008 to 2025 are available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52013XC0802(04).

[4] https://www.fda.gov/news-events/speeches-fda-officials/capturing-benefits-competition-patients-0307

3.2.2 Data

Studying learning-by-doing in the production of biologic drugs requires information on drug characteristics and production process, including product-level output and manufacturing costs at each location. This information is not commonly available for pharmaceuticals. As a result, we overcome this data limitation by combining four different types of information: (*i*) biologic drug registries, (*ii*) product-level quarterly sales data by country from 2002 to 2022, (*iii*) product-level production facilities, and (*iv*) product-level historical data on production process variations post-market entry.

Biologic and biosimilar drugs Biologic drugs approved in the US are registered in the FDA Purple Book. In the EU, they are included in the list of medicines approved by the European Medicines Agency (EMA).[5] The dataset provides information on marketing authorization within the EU, including whether the product was approved as a biosimilar. Hence, among biologic drugs, we identify biosimilar products from the EMA registry. The biosimilar status may differ depending on the geographic region. For instance, Basaglar was approved as a biosimilar by the EMA in Europe and as a biological product by the US FDA. As highlighted by Scott Morton et al. (2018), the biosimilar approval pathway was first introduced in the EU. Throughout this paper, we use the EMA product application to identify biosimilars from traditional biologic products.

Drug quantities and prices To measure global drug production, we rely on country-level sales data from IMS Health which became IQVIA in 2016. The dataset provides quarterly revenues and quantities sold for each drug from 2002 to 2022 at the country level. Our dataset includes information for 64 countries, including 25 covering the entire sample period. The list of countries available is provided in Table 3.6. Table 3.5 focuses on countries available from 2002 to 2022. Although not exhaustive, the set of countries included in our dataset represents a significant share of global pharmaceutical sales.[6] As a result, for countries in Table 3.5, we observe sales data since market entry for biologic drugs approved after 2001.[7] The volume data covers quantities consumed in each country. The product-level revenues allows us to recover drug prices. In the descriptive analysis, we use standard units to compute quantities and

[5] https://www.ema.europa.eu/en/medicines/download-medicine-data

[6] The set of countries available varies between 2002-2010 and 2010-2022. Sales in Europe and North America are included over the whole sample period. Partial data for Asia and Latin America are available until 2010.

[7] Biologic drugs are relatively recent. Products approved since 2002 account for two-thirds of biologic medicines licensed in the US in 2023 (Figure 3.5).

per-unit price at the country-product-quarter level to ensure comparability between the forms and strengths of a particular product. In the demand model estimation, we however use the quantities in Defined Daily Dose (DDD) which is what determines the quantity a patient is supposed to use per fay of treatment.

Production facilities We recover the location of production facilities for biologic drugs and biosimilars from publicly available sources. For biologic and biosimilar drugs approved in the EU, the manufacturers of the active substance and those responsible for batch release are listed in the Summary of Product Characteristics, available on the EMA website. This data source presents three main caveats: it only provides partial information about production facilities (i) for a given drug, (ii) across markets, and (iii) over time. Indeed, the manufacturing process may rely on sites other than the ones responsible for the active substance and batch release. Hence, we observe only a subset of factories involved in the production process.[8] Moreover, pharmaceutical companies may produce biologic drugs in various locations worldwide, such that manufacturers approved in the EU represent only a subset of production facilities. To overcome this, we complement the location approved by the EMA with information about manufacturing sites approved by the FDA for the US market. In particular, approval letters provide information on sites approved for the manufacturing of the active substance and locations where the formulated drug product will be manufactured, filled, labeled, and packaged. This information is often confidential and therefore unavailable. For each drug, the list of manufacturing sites is obtained by merging the list of facilities approved by the EMA and FDA.[9] Last, the information available on the EMA website provides a snapshot of approved manufacturers as of 2025. While changes in production facilities are registered in the "Procedural steps taken and scientific information after the authorization", the location of previous factories remains unavailable.[10] These caveats must be kept in mind when analyzing the data. Table 3.8 provides summary statistics about manufacturing sites for one particular class of biologic drugs, TNF-α inhibitors. The average number of active substance manufacturing sites is 1.6 per product, suggesting that the production of the active substance is relatively concentrated. The number of facilities responsible for batch release remains low, at 1.85 on average.

[8] Such facilities include formulation and filling sites (if they differ from batch release sites), secondary packaging or labeling sites, testing-only facilities (e.g. for sterility, stability, potency), and Cold chain logistics centers.

[9] Concerns about manufacturing facilities not registered in the EU nor US is mitigated by the fact that, in the case of TNF-α inhibitors, 60% of active substance manufacturers approved by the EMA are located outside Europe.

[10] In future iterations, we plan to track changes in production facilities from the partial past information available on the Internet archive.

Changes in the production process We recover variation in the production process of biologic drugs from the "Procedural steps taken and scientific information after the authorization" documents available on the EMA website. These documents summarize all changes that were implemented in the manufacturing of each biologic drug since market approval. They are informative of the occurrence of production process updates, the type of changes that are deployed, and when these changes were notified to the regulatory agency. In particular, in the EU, new guidelines were published by the EMA in 2013 regarding the procedures outlined in the 2008 regulation. These guidelines provide application codes for each scope of changes that are executed. They classify variations into three main categories: (A.) Administrative changes (B.) Quality changes and (C.) Safety, Efficacy, and Pharmacovigilance Updates. We consider changes in paragraphs B.I. ('Active Substance') and B.II. ('Finished Product') to affect 'product' manufacturing. The document also provides information about the corresponding notification requirements for each variation ("Do and tell", "Tell, wait and do", and prior approval), which is informative of the adjustments flexibility. As mentioned above, "Do and tell" changes can be applied immediately, while the other two categories require notification or authorization before implementation. The discrepancy between the notification date and actual deployment date (either before or after the notification) prevents inferring the exact date of the changes without proprietary data.[11] Changes carried out before 2013 are registered, but for the vast majority, cannot be matched with the corresponding procedure.

To build our final dataset, the first three data sources are linked as follows. First, biologic drugs are extracted from the sales data based on molecule names. The biosimilar status is inferred based on the EMA list. The first biosimilar was approved in the EU in April 2006; therefore, our data covers the introduction of any biosimilar until 2022. The descriptive analysis presented in section 3.3 focuses on all biologic drugs facing biosimilar competition. The final sample includes 13 molecules and 57 biosimilars, presented in Table 3.7. The number of biosimilars available per molecule ranges from 1 to 8. In the structural model presented in section 3.4, we focus on a particular drug class, TNF-α inhibitors, corresponding to the Anatomical Therapeutic Chemical (ATC) class L04AB. To link sales data with manufacturing sites, we allocate sales in each country to the closest production facilities using the distance

[11] In particular, we observe the notification/opinion date and commission date (when applicable). "Do and tell" variations must be notified at the latest within 12 months from the implementation date. Only a subset of "Do and tell" variations must be notified immediately (Type IA_{IN}). In other cases (such as "Tell, wait and do" and prior authorization), deployment must occur after the notification and/or decision, when applicable.

between the main cities of the sales and manufacturing countries. Each product sold in each country is associated with its closest active substance and batch release manufacturers. Note that, as most active substances for a given drug are produced in one location, our measure of volume produced at a given active substance manufacturer often coincides with our measure of global volume produced.

3.3 Empirical evidence

This section presents motivating evidence that learning-by-doing may arise when producing biologic drugs and biosimilars. Learning-by-doing captures the idea that firms become more cost-efficient as they accumulate knowledge producing a product. This phenomenon translates into marginal costs of production decreasing with experience in that product. This feature can be observed in product-level production cost data, a piece of information often unavailable in the context of pharmaceuticals. On the other hand, product market data provides information on product-level prices and quantities. Assuming a simplified static Bertrand-Nash pricing game, observed prices depend on marginal costs. In the presence of learning-by-going, marginal costs decrease with experience such that, *keeping everything else constant*, prices should be negatively correlated with experience levels. One of the key elements is the definition and measurement of *experience*. Learning-by-doing suggests that cost efficiencies arise not only as a function of time but are driven by *how much* you produced in the past. Unlike learning happening only through time, the intensity of production affects the learning rate. In this case, experience becomes a strategic variable the firm can manipulate, and firms may decide to set prices dynamically such that observed prices deviate from the static Bertrand-Nash equilibrium outcome.

This paragraph documents learning in biologic drugs manufacturing through the correlation between drug prices and experience levels. We consider two alternative measures of experience: (i) the age of the drug and (ii) the volume of past drug sales, under the assumption that learning in production experience increases with past production volumes. Current prices are expected to be negatively correlated with experience, as efficiency gains reduce the marginal costs of the drug. We consider the price of drug j of type s (branded or biosimilar) corresponding to molecule m, in quarter t and country r,

$$\log(p_{jtr}) = \alpha_{s(j)} \log\left(\sum_{r} \sum_{t'=0}^{t-1} q_{jt'r}\right) + \gamma \log(age_{jt}) + \beta_1 N_{mtr} + \beta_2 \log Q_{-jtr} + \delta_j + \delta_{s(j),r} + \delta_t + \varepsilon_{jtr} \tag{3.1}$$

where q_{jtr} corresponds to sales volume at t, age_{jt} is the age of drug j at time t, N_{mtr} the number of product from molecule m available in country r in quarter t and Q_{-jtr} the total past sales of products other than j but with the same molecule $Q_{-jtr} = \sum_{t'=0}^{t-1} \sum_r \sum_{j' \in \mathcal{J}_m \setminus j} q_{j't'r}$. We include fixed effects that capture heterogeneity across drugs, countries, and time. We allow for the correlation with past sales to differ for branded and biosimilar products, and, in subsequent specifications, for the correlation between the age of the drug and its price to differ for acute vs chronic treatments.[12] The results are presented in Table 3.1. The stock of past quantities is negatively correlated with the product price, while the product age appears to be positively correlated with price once we control for past quantities. The negative correlation is primarily driven by biosimilars, while the positive correlation between age and prices is mostly observed for treatments of chronic conditions. A negative correlation is also observed between past volumes and current prices when considering only TNF-α inhibitors, a drug class that we focus on in the second part of this paper.

This section suggests a negative correlation between past sales and current prices. This correlation is not solely driven by the heterogeneity across products or changes in a particular molecule over time. We also rule out that this correlation is solely driven by drug aging or sales from other manufacturers. More precisely, when considering a country with a small impact on total sales, past sales in other countries remain negatively correlated with current prices. This correlation does not arise in regions with regulated prices, such as France, Germany, or the UK (as shown in Table 3.10 in appendix). The next section quantifies learning-by-doing in the production process of biologic drugs, specifically in the context of TNF-α inhibitors, by combining sales data with a model of drug demand and pricing.

3.4 Biologic demand and supply model: TNF-α inhibitors

This section aims to quantify learning-by-doing in the context of TNF-α inhibitors, overcoming the lack of production data. We estimate product and time-specific marginal costs by combining a model of biologic drug demand and pricing with sales data. We then rely on a marginal cost model to measure the contribution of learning-by-doing to cost efficiencies.

[12] Demand for chronic treatments is subject to inertia, such that drug manufacturers may have incentives to set prices differently for drugs treating acute and chronic conditions over the product lifetime.

Table 3.1: Determinants of prices with respect to past experience

Countries	(1) All All drugs	(2) All All drugs	(3) All All drugs	(4) All TNF-α	(5) Sweden TNF-α	(6) Sweden TNF-α
α_s non biosimilar	-0.334***	-0.00557	0.000223	0.0212***	0.00716	
	(0.00754)	(0.00677)	(0.00710)	(0.00671)	(0.0206)	
α_s biosimilar	-0.611***	-0.0261***	-0.0188***	-0.0352***	-0.0427*	
	(0.00867)	(0.00542)	(0.00642)	(0.00528)	(0.0230)	
Log(age) (γ)	1.227***			0.131***	-0.0968	-0.0945
	(0.0215)			(0.0209)	(0.0861)	(0.0855)
γ acute		-0.000280	-0.00635			
		(0.0259)	(0.0253)			
γ chronic		0.0955***	0.0746***			
		(0.0214)	(0.0224)			
Nb products (β_1)	-0.115***	-0.0332***	-0.0392***	-0.0645***	-0.0709***	-0.0706***
	(0.00721)	(0.00355)	(0.00366)	(0.00383)	(0.0143)	(0.0142)
β_2			0.00824*			
			(0.00472)			
Exchange rate					-0.161***	-0.161***
					(0.0137)	(0.0137)
α_s non biosimilar (Except Sweden)						0.00672
						(0.0205)
α_s biosimilar (Except Sweden)						-0.0434*
						(0.0230)
Time FE	Yes	Yes	Yes		Trend	Trend
Year × Country FE		Yes	Yes	Trend		
Product FE		Yes	Yes	Yes	Yes	Yes
Biosimilar × Country FE	Yes	Yes	Yes	Yes		
Molecule × Year FE	No	Yes	Yes	Trend	Trend	Trend
Constant	4.646***	5.508***	4.900***	6.527	-31.12*	-30.78*
	(0.0558)	(0.0745)	(0.0562)	(7.209)	(17.94)	(17.84)
Observations	32,519	32,519	22,809	9,267	446	446
R-squared	0.371	0.949	0.936	0.862	0.948	0.949

Note: Robust standard errors in parenthesis.

3.4.1 TNF-α inhibitors

Among the molecules presented in Table 3.7 in appendix, we restrict our attention to TNF-α inhibitors when estimating marginal costs. These drugs are identified by the L04AB ATC level 4 class. We focus on this class of drugs for several reasons. First, biosimilars competition started to arise from 2013, and three out of five molecules are facing biosimilar competition by 2022. The sales data cover all sales since biosimilar entry across 34 countries. These countries represent, on average, 70% of the countries in which a given biosimilar product is available, accounting for 2/3 of the total population (Table 3.9). Second, given the high price of the branded version, the stakes in this market are high both for drug manufacturers and healthcare payers. Between three and eight different biosimilar products are available, and their characteristics

vary slightly from one biosimilar to another (see Table 3.12). Drug prices vary across products and over time in some countries, and biosimilars represent a significant share of volumes sold over the sample period, ranging from 6.7 to 14%. Third, the closest substitutes are easily identifiable as they belong to the same ATC4 class, and there have been no changes in pharmacy substitution rules over time. Lastly, quantities in Defined Daily Doses (DDD) are available in the sales data, allowing for the consistent measurement of volumes and market shares across molecules.

TNF-α inhibitors are injectables, mostly used to treat chronic conditions, including rheumatoid arthritis, spondylarthritis, plaque psoriasis, and chronic inflammatory bowel diseases. The frequency of injection varies from one molecule-condition combination to another. It can be performed at home (using disposable injection devices such as pens) or supervised by a care provider at the hospital. Molecule-specific information is presented in Table 3.11. As mentioned above, there are differences between biosimilars and their reference product (see Table 3.12). While biosimilars are equivalent to the reference product, they are not exact copies. Biosimilars follow strict FDA or EMA guidelines to ensure they work the same and are just as safe as brand name biologic medicines, they are very close in structure and function to an existing brand name biologic medicine. A biosimilar behaves in much the same way as its brand name biologic, is considered as safe and effective as the biologic reference product, is made with the same types of natural sources, is given the same way, provides the same treatment benefits, has the same potential side effects, has the same strength and dosage. However, while a generic drug has the same chemical makeup, biologic medicines come from natural, living sources and can't be copied exactly because natural sources are very complex and their environments can change.

The production of TNF-α inhibitors is rather concentrated, with on average 1.6 manufacturing sites for the active substance per product and 1.85 sites responsible for batch release (Table 3.8). Facilities producing the active substance are primarily located in Europe (40.6 %), Asia (40.6%), and North America (18.75%), while facilities responsible for batch release are mostly located in Europe.

The manufacturing process of these products evolved after market approval. Based on the EMA registries, there are 81 applications per product on average since market approval and 67.5 changes in the production process (Table 3.14). One application can group several variations.

Nearly half of the changes (46.6%) are directly related to the drug manufacturing. These variations affect either the active substance (41.5%) or the final product (58.5%) step. More than 60% of 'Product' variations are related to the drug's manufacturing (excluding changes in production sites) and testing (Table 3.13). Overall variations are frequent, with on average, 1.7 applications per quarter by product since market entry (Table 3.14). The average time between two 'Product' variation notifications is 87 days, suggesting that the process is updated quite regularly (Figure 3.7). These notifications appear slightly more frequently right after market entry and decrease over time (Figure 3.8). 36% of these changes are implemented immediately, before the notification date ("Do and tell").

3.4.2 Biologic drug supply with learning-by-doing

Production

We consider the production of drug j, offered in quarter t, in country $r \in R_{jt}$, by manufacturer m. Each drug is produced by one manufacturer. The production of a given drug can be scattered across different factories, $f \in \mathcal{F}_m$ but we assume that one unit is produced at the same factory throughout its process. Hence, we integrate manufacturing sites producing the active substance with those responsible for the final formulation, packaging and batch release. We further assume that units supplied in a given country are produced in the same factory, $f(r)$, the closest site to its capital city. To produce one additional unit of drug j in quarter t and factory f, each firm is facing the following marginal cost function:

$$c_{jt}^{f(r)}(e_{jft}, Q_{jft}) = \kappa_{jft} e_{jft}^{\eta} Q_{jft}^{\gamma} \tag{3.2}$$

where e_{jft} represents the experience of drug j's manufacturer in factory f accumulated until the beginning of period t and Q_{jft} is the total volume produced at t. This cost function captures two important features of drug manufacturing. γ measures potential economies of scale where the marginal cost decrease with the number of units produced in a given period. $\eta \neq 0$ ensures that the current marginal cost depends on experience. When $\eta < 0$, past production decreases current marginal costs, characterizing learning-by-doing. The specification follows Besanko et al. (2010). The experience of j's manufacturer at factory f in time t is

$$e_{jft} = \sum_{f' \in \mathcal{F}_m} w_{f'} \left(\sum_{\tau=0}^{t-1} \sum_{r \in R_{j\tau}, f(r)=f'} s_{jr\tau}(p_{jr\tau}) M_{r\tau} \right) \tag{3.3}$$

where $\mathcal{F}_m$ is the set of manufacturing sites of drug j and $w_f \in [0,1]$ the weight that this manufacturing site has on the firm experience, $s_{jrt}(p_{jrt})$ corresponds to the market share of product j in country r and quarter t and M_{rt} to r's market size. The manufacturer's experience in factory f depends on past sales in all countries where the product is available up to the previous period, hence it is not only a matter of time passing. This specification assumes that volume produced in the facility, f, at t can be approximated by the sum of the sales in countries the factory is producing for, $r \in R_{jt'}$, such that $f(r) = f$, at t. There is no delay between the timing of production and sales. As production at $t-1$ directly affects marginal cost at t, adjustments to the production process are implemented rapidly. The experience in factory f is a weighted sum of past sales in any factory in which drug j is produced, allowing manufacturers to learn across factories. e_{jft} only depends on past sales of the same drug j, assuming there are no experience spillovers across manufacturers. The expertise acquired through learning-by-doing remains proprietary. While the current specification implies decreasing marginal return to experience, unlike Cabral and Riordan (1994) and Besanko et al. (2010), we assume bottomless learning. We also abstract away from organizational forgetting (Benkard (2000), Besanko et al. (2010), Thompson (2010)). Hence, the factory's state, represented by the weighted sum of past sales, continues to increase over time. The per-period profits for firm j are given by

$$\pi_{jt} = \sum_{r \in R_{jt}} \Big(p_{jrt} - c_{jt}^{f(r)}(e_{jt}, Q_{jft})\Big) s_{jrt}(p_{jrt}) M_{rt} \tag{3.4}$$

leading to the following discounted flow of profits

$$\Pi_j = \sum_{t=0}^{T} \beta^t \sum_{r \in R_{jt}} \Big(p_{jrt} - c_{jt}^{f(r)}(e_{jt}, Q_{jft})\Big) s_{jrt}(p_{jrt}) M_{rt} \tag{3.5}$$

Drug pricing

Drug manufacturers are setting their prices to maximize their discounted flow of profits given by Equation 3.5. When producers are subject to learning-by-doing, equilibrium prices may deviate from the static Bertrand-Nash equilibrium prices for two reasons. First, manufacturers are using price cuts to move down their learning curve (and decrease their future marginal costs). Second, they use their price to prevent their competitors from winning and moving down their learning curve. As a result, we deviate from Khmelnitskaya et al. (2024) who assumes static pricing when estimating scale and scope economies fro product-market data. Moreover, in the

context of multi-market products, spillovers across markets arise as well: the price in country r affects future marginal costs which enter the objective function when setting price in r'. Decentralized pricing decision at the country level will fail to take into account the impact of current prices on marginal costs beyond country r and may lead to slower learning, underestimating the returns to experience.

The objective of this paragraph is to illustrate how equilibrium prices deviate from the static Bertrand-Nash prices when the manufacturer internalizes learning-by-doing and the impact of pricing decentralization. Let's consider a simple example of a manufacturer m producing product j for two countries, $r \in \{a, b\}$, with market size M_a and M_b, over two periods, $t = \{1, 2\}$. The units are produced in the same factory, f. The net present value of the firm is given by $\pi_{j1} + \beta\pi_{j2}$. In the last period ($t = 2$), the price in country r is set according to Bertrand-Nash pricing:

$$p_{jr2} = c_{j2}(e_{j2}, Q_{jf2}) - \frac{(1 - \partial c_{j2}/\partial p_{jr2})s_{jr2}}{\partial s_{jr2}/\partial p_{jr2}} + \frac{\partial c_{j2}/\partial p_{jr2} \times s_{jr'2}M_{r'}}{\partial s_{jr2}/\partial p_{jr2}M_r} \tag{3.6}$$

where $r' = \{a, b\}\backslash r$, $e_{j2} = s_{ja1}(p_{ja1})M_a + s_{jb1}(p_{jb1})M_b$ is the total quantities sold in the previous period and $\partial c_{j2}/\partial p_{jr2} \neq 0$ due to potential economies of scale. In period 1, the expected net present value of future profits is

$$\begin{aligned} V_{j1} &= \pi_{j1} + \beta\mathcal{E}(V_{j2}) \\ &= \Big(p_{ja1} - c_{j1}(e_{j1}, Q_{jf1})\Big)s_{ja1}(p_{ja1})M_a + \Big(p_{jb1} - c_{j1}(e_{j1}, Q_{jf1})\Big)s_{jb1}(p_{jb1})M_b + \beta\mathcal{E}(V_{j2}) \end{aligned}$$

where $\mathcal{E}(.)$ captures firm's f uncertainty on unobserved demand and costs shocks. From the first order condition of the firm, the equilibrium prices with learning-by-doing satisfy

$$p^{LBD}_{ja1} = c_{j1}(e_{j1}, Q_{jf1}) - \frac{(1 - \partial c_{j1}/\partial p_{jr1})s_{ja1}}{\partial s_{ja1}/\partial p_{ja1}} + \frac{\partial c_{j1}/\partial p_{jr1} \times s_{jb1}M_b}{\partial s_{ja1}/\partial p_{ja1}M_a} - \frac{\beta}{\partial s_{ja1}/\partial p_{ja1}M_a}\frac{\partial \mathcal{E}(V_{j2})}{\partial p_{ja1}} \tag{3.7}$$

If price settings are independents across countries (in "autarky"), prices at $t = 2$ satisfy

$$p^{A}_{jr2} = c_{j2}(e_{j2}, Q_{jf2}) - \frac{(1 - \partial c_{j2}/\partial p_{jr2})s_{jr2}}{\partial s_{jr2}/\partial p_{jr2}} \tag{3.8}$$

At $t = 1$, the maximization of the expected net present value in country a leads to

$$p^A_{ja1} = c_{j1}(e_{j1}, Q_{jf1}) - \frac{(1 - \partial c_{j1}/\partial p_{ja1})s_{ja1}}{\partial s_{ja1}(p_{ja1})/\partial p_{ja1}} - \frac{\beta}{\partial s_{ja1}/\partial p_{ja1} M_a} \frac{\partial \mathcal{E}(V^a_{j2})}{\partial p_{ja1}} \tag{3.9}$$

where $V^a_{jt} = \Big(p_{jat} - c_{jt}(e_{jt}, Q_{jt})\Big)s_{jat}(p_{jat})M_a$. In equations (3.7) and (3.9), $\mathcal{E}(V^a_{j2})$ is particularly complex due to strategic interactions between drug manufacturers. In what follows, we consider the simplified case without strategic interactions to highlight the differences with myopic Bertrand-Nash pricing.

We compare equilibrium prices from forward looking firms without strategic interactions to the myopic Bertrand-Nash pricing in the first period. The main difference between the two models depends on how firm j expects prices in the first period to affect its profits in second period, $\partial \mathcal{E}(V_{j2})/\partial p_{ja1}$. In this simplified version, we also abstract away from economies of scale, setting $\gamma = 0$ as we specify $c^{f(r)}_{jt}(e_{jft}, Q_{jft}) = \kappa_{jft} e^{\eta}_{jft} Q^{\gamma}_{jft}$ (Equation 3.2). With these simplifications, we have

$$\frac{\partial \mathcal{E}(V_{j2})}{\partial p_{ja1}} = -\Bigg(M_a s_{ja2} + M_b s_{jb2}\Bigg) \frac{\partial c_{j2}(e_{j2}, Q_{jf2})}{\partial p_{ja1}}$$

and

$$\frac{\partial \mathcal{E}(V^a_{j2})}{\partial p_{ja1}} = -M_a s_{ja2} \frac{\partial c_{j2}(e_{j2}, Q_{jf2})}{\partial p_{ja1}}$$

This implies that,

$$\frac{\partial e_2}{\partial p_{ja1}} = \frac{\partial s_{ja1}(p_{ja1})M_a + s_{jb1}(p_{jb1})M_b}{\partial p_{ja1}} < 0$$

and, in the presence of learning-by-doing, $\eta < 0$, we obtain that:

$$\frac{\partial c_{j2}(e_{j2}, Q_{jf2})}{\partial p_{ja1}} = \kappa\eta \frac{\partial e_2}{\partial p_{ja1}}(e_2^{\eta-1}) > 0 \; ; \; \frac{\partial \mathcal{E}(V_{j2})}{\partial p_{ja1}} < 0 \; ; \text{ and } \frac{\partial \mathcal{E}(V^a_{j2})}{\partial p_{ja1}} < 0$$

When firms are facing learning-by-doing, past sales are affecting current prices through marginal costs reduction and, forward-looking firms will price differently than myopic ones: $p^{BN}_{ja1} > p^A_{ja1} > p^{LBD}_{ja1}$. Learning-by-doing leads to lower price upon entry to acquire experience ($p^{BN}_{ja1} > p^{LBD}_{ja1}$). Even without economies of scale ($\gamma = 0$), independent price negotiations across countries lead to higher prices ($p^A_{ja1} > p^{LBD}_{ja1}$) as the firm does not internalize the dynamic externalities. Autarkic prices are still lower than Bertrand-Nash prices p^{BN}_{ja1} and the difference between myopic Bertrand Nash and Autarkic prices decreases with the market size, M_a.

3.4.3 Demand for TNF-α inhibitors

The demand model aims to recover substitution patterns between branded drugs, their biosimilar counterparts, and the outside option. In particular, the substitution to the outside option is key in our setup to allow for more patients to be treated with TNF-α inhibitors thanks to biosimilar competition.

In a given quarter, t, and country r, an individual i, diagnosed with Rheumatoid Arthritis, Spondylarthritis, Plaque Psoriasis or Chronic Inflammatory Bowel diseases, together with her physician decides over the treatments $j \in \mathcal{J}_{rt} \cup 0$ where $\mathcal{J}_{rt}$ includes TNF-α inhibitors available in quarter t and country r. The outside option, indexed by 0, captures the substitution toward treatments other than TNF-α inhibitors, including no treatment.The decision utility for patient i in country r from taking drug j in quarter t is

$$U_{ijrt} = X_{jrt}\beta + \alpha p_{jrt} + \xi_{jrt} + \nu_{ijrt} \tag{3.10}$$

where X_{jrt} corresponds to drugs characteristics, p_{jrt} the price per DDD in national currency units, and ν_{ijrt} corresponds to the idiosyncratic taste shock. X_{jrt} includes drug fixed effects, and an indicator variable for parallel sales and the logarithm of the product's age in quarter t. Indeed, TNF-α inhibitors are drugs used to treat chronic conditions, hence demand is likely to be subject to inertia. We account for this feature by allowing the preference for a given product to vary over its lifetime.[13] We further assume that the decision utility can be represented by a nested logit model such that the substitution patterns within and across groups differ. More precisely, we separate the inside and outside goods in different nests, such that $\nu_{ijrt} = \zeta_{ig(j)} + (1-\sigma)\varepsilon_{ijrt}$ where ε_{ijrt} is Type I EV distributed. σ measures the strength of the correlation in unobserved preference for products within a nest. The indirect utility for the outside good is normalized to $\zeta_{i0} + (1-\sigma)\varepsilon_{i0rt}$.

The market share for the (inside) product j in country r can be written as

$$s_{jrt}(\mathbf{p}_{rt}) = \frac{\exp\left(\frac{X_{jrt}\beta + \alpha p_{jrt} + \xi_{jrt}}{1-\sigma}\right)}{\sum_{\forall k \in \mathcal{J}_t} \exp\left(\frac{X_{krt}\beta + \alpha p_{krt} + \xi_{krt}}{1-\sigma}\right)} \frac{\exp(I_g)}{1+\exp(I_g)} \tag{3.11}$$

[13] MacKay and Remer (2024) provide an empirical model to accommodate inertia on the demand side using market-level data.

where $\exp(I_g) = (1-\sigma)\log\left(\sum_{\forall k \in \mathcal{J}_t} \exp\left(\frac{X_{krt}\beta + \alpha p_{krt} + \xi_{krt}}{(1-\sigma)}\right)\right)$.

3.5 Estimation

Section 3.4 highlights the dynamic considerations at stake in the context of learning-by-doing. In particular, drug manufacturers may internalize future learning-by-doing when setting their current prices, leading to forward-looking behavior. This feature complicates marginal cost estimation as it requires solving for the dynamic game played by drug manufacturers. Moreover, in some countries such as France or the UK, biosimilar prices are tightly regulated. Any biosimilar for the same molecule is sold at the same price. The lack of price variation across products complicates the estimation of product-specific marginal costs.

We rely on the demand and pricing of TNF-α inhibitors in Sweden to overcome these challenges and estimate marginal costs. First, in Sweden, prices for prescription drugs reimbursed by the national health service are submitted by pharmaceutical companies to the Swedish Reimbursement Authority (TLV). While generic prices are closely tied to the price of their branded version, there is no link between the prices of the original biologic drugs and their biosimilar counterparts, or across biosimilars (Pontén et al. (2017)). As a result, prices for biosimilars of the same reference drug vary across products and across time. Second, Sweden is a small market. Between 2010 and 2022, it represents only 1.5% of total product sales for TNF-α inhibitors, and only up to 8% for a given drug (with an average of 3%). Hence, Swedish volumes contribute marginally to learning-by-doing, compared to volumes sold in France or Germany, for example. Third, we assume that drug manufacturers set prices for Sweden independently of other countries, abstracting from External Reference Pricing (Maini and Pammolli (2023)).[14] As highlighted in section 3.4.2, the difference between static Bertrand-Nash and autarkic dynamic pricing is small in small markets. As a result, we assume static Bertrand-Nash pricing in Sweden and recover the marginal costs of a given drug j in quarter t from the first order condition of the firm.[15] As the first order condition of the firm depends on the price elasticity

[14] While prices in Sweden are not subject to External Reference Pricing (ERP), Swedish prices are used as a reference in up to 13 countries. According to Rémuzat et al. (2015), on average, countries using Swedish prices for ERP rely on prices from 20 countries, ranging from 4 (Cyprus, Iceland) to 31 areas (Hungary, Poland). Maini and Pammolli (2023) suggest fewer countries are using Swedish prices as a reference.

[15] Between 2016 and 2021, Managed Entry Agreements were signed between TNF-α manufacturers and the Swedish Reimbursement Authority (TLV). The product-level rebate rate remains confidential. Yet, publicly available data provide information about the agreement's date and products, as well as the value of the total rebate per drug class per year for three years. In future iterations, we will exploit this information to estimate

of demand, we recover price sensitivity thanks to demand estimation.

3.5.1 Demand

The demand for TNF-α inhibitors in Sweden follows the nested logit model presented in section 3.4. The model is estimated from aggregate data using the inversion from Berry (1994). Prices are expressed in Swedish krona (SEK) per DDD. The market size is computed from the number of patients diagnosed each year for Rheumatoid Arthritis, Spondylarthritis, Plaque Psoriasis and Chronic Inflammatory Bowel diseases in inpatient or specialized care multiplied by 90 days to recover the quarterly market size in DDD.[16] The estimating equation is thus:

$$\log s_{jrt} - \log s_{0rt} = \alpha p_{jrt} + X_{jrt}\beta + \sigma \log s^0_{jrt} + \xi_{jrt} \tag{3.12}$$

In equation (3.12), the within-nest market shares and prices are endogenous, hence potentially correlated with the structural error term, ξ_{jrt}. The identification of the within-nest unobserved preference correlation, σ, and the price sensitivity parameter, α is obtained by relying on instrumental variables, shifting prices, and the within-nest market share independently of ξ_{jrt}. We rely on the competition instrumental variables that are standard in the literature, capturing the change in the product space over time due to the entry of new drugs. We use the number of products produced by other drug manufacturers, the number of competing products within a molecule, and the number of products that share the same citrate-free and latex-free characteristics.[17] The final model is estimated by 2SLS. Table 3.2 presents the results. Columns 1 and 2 present the estimation of a multinomial logit model. Column 3 the nested logit. Column 4 includes product fixed effects in the nested logit and corresponds to our preferred specification. In columns 3 and 4, the nest parameter, σ, is statistically significant, such that we can reject the multinomial logit specification. In column 4, the estimated nest parameter is 0.8, suggesting a strong correlation in the unobserved preferences within inside goods. The price sensitivity estimate is -0.00731 while the mean price is 215.52 SEK/DDD.[18]

rebate rates and use rebated prices instead of list prices.

[16] We use annual number of individuals diagnosed with one of these disease, corresponding to the following International Classification of Diseases (ICD) 10: Rheumatoid Arthritis (M05, M06), Spondylarthritis (M45, M46), Psoriasis (L40), Chronic Inflammatory Bowel (K50, K51). Source: https://sdb.socialstyrelsen.se/if_par/val_eng.aspx

[17] Product characteristics are presented in Table 3.12. Citrate is an excipient commonly used that can be responsible for pain during the drug's injection. Latex can generate allergic reactions.

[18] It corresponds to 26.73 USD/DDD.

Table 3.2: Demand estimation

	(1) MNL OLS	(2) MNL IV	(3) NL IV	(4) NL FE IV
Price (α)	-0.00616***	-0.00960***	-0.00751***	-0.00731***
	(0.00171)	(0.00270)	(0.00116)	(0.00154)
s^0_{jrt} (σ)			0.645***	0.803***
			(0.0651)	(0.0199)
Branded dummy	1.908***	2.573***	0.930***	
	(0.453)	(0.606)	(0.323)	
citrate free dummy	1.329***	1.318***	0.589***	
	(0.287)	(0.286)	(0.130)	
latex free dummy	-0.214	-0.381	-0.146	
	(0.288)	(0.304)	(0.122)	
parallel dummy	-3.135***	-3.130***	-1.120***	-0.614***
	(0.167)	(0.166)	(0.212)	(0.0806)
age	0.405***	0.395***	0.168***	
	(0.0534)	(0.0534)	(0.0299)	
age^2	-0.0219***	-0.0225***	-0.00817***	
	(0.00228)	(0.00229)	(0.00173)	
log(*age*)				0.233***
				(0.0605)
Molecule dummies				
Certolizumab pegol	-0.815**	-0.830**	-0.182	
	(0.346)	(0.344)	(0.146)	
Etanercept	0.415	0.595*	0.0714	
	(0.305)	(0.323)	(0.140)	
Golimumab	-0.851**	-0.833**	-0.145	
	(0.354)	(0.352)	(0.149)	
Infliximab	0.369	0.219	-0.429***	
	(0.354)	(0.363)	(0.147)	
Product FE				Yes
Constant	-5.421***	-4.881***	-1.250***	0.552
	(0.374)	(0.496)	(0.371)	(0.713)
Observations	644	644	644	632
R-squared	0.432			
Weak IV Fstat		138	11.47	19.27

Notes: Column 1 presents estimates from a multinomial logit model that does not account for price endogeneity. Column 2 introduces the instruments to correct for price endogeneity. Column 4 allows for unobserved preference correlation within inside goods. Column 4 includes product fixed effects. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

3.5.2 Pricing

The features of the Swedish context enable us to recover the marginal cost of drug j in quarter t from the first-order condition of a static Bertrand-Nash pricing game between single-product firms. For $r = Sweden$, denoting $c_{jt}^{f(r)}$ the marginal cost in factory $f(r)$ of drug j in quarter t, the equilibrium prices, p_{jrt}, are satisfying

$$p_{jrt} = c_{jt}^{f(r)} - \frac{\tilde{s}_{jrt}}{\partial \tilde{s}_{jrt}/\partial p_{jrt}} \tag{3.13}$$

where $\tilde{s}_{jrt}$ corresponds to the expected market share in Sweden. We compute $-\frac{\tilde{s}_{jrt}}{\partial \tilde{s}_{jrt}/\partial p_{jrt}}$ using the estimates presented in Table 3.2 column 4 and recover product-quarter specific marginal costs, $c_{jt}^{f(r)}$. Table 3.3 summarizes the results. The average markup is 25.7%, displaying little heterogeneity across molecules. In two out of three cases, the markup is, on average, higher for biosimilars compared to their reference products. Within each product, marginal costs vary significantly over time. The within-product standard deviation of marginal cost over time represents 21.8% of the average price, ranging from 5 to 32.3% on average for biosimilars.

Table 3.3: Marginal costs estimates

Molecule	Biosimilar	Marginal cost Mean	Marginal cost SD	Markup Mean	Marginal cost SD within
Adalimumab	0	238.8	47.37	0.232	0.251
	1	63.96	30.89	0.348	0.323
Certolizumab Pegol	0	275.2	.	0.115	0.0994
Etanercept	0	280.0	.	0.0949	0.151
	1	126.9	16.30	0.325	0.524
Golimumab	0	247.1	.	0.212	0.159
Infliximab	0	114.8	.	0.381	0.225
	1	91.29	21.54	0.249	0.0494
All	0	216.4	64.13	0.236	0.209
	1	92.93	29.33	0.292	0.232
All		169.7	84.57	0.257	0.218

Notes: Column 1 presents the average marginal cost across products, weighted by the total sales of each product. Column 2 presents the standard deviation of marginal costs across products. Column 3 computes the average markup across products. Column 4 presents the average of the within-product standard deviation of marginal cost, expressed as a percentage of the average price.

3.5.3 Cost function estimation

In this section, we rely on the marginal cost estimates presented in Table 3.3 to estimate the biologic drug cost function. In particular, the objective of this section is to estimate the learning-by-doing and economies of scale parameters, η and γ. Given the marginal cost function presented in paragraph 3.4.2,

$$\log\left(c_{jt}^{f(r)}(e_{jft}, Q_{jft})\right) = \log(\kappa_{jft}) + \eta\log(e_{jft}) + \gamma\log(Q_{jft}) \tag{3.14}$$

We estimate an empirical counterpart of expression (3.14), given by:

$$\log(\tilde{c}_{jt}^{f(r)}) = \eta\log(e_{jft}) + \gamma\log(Q_{jft}) + \lambda_j + \delta Time_t + \varepsilon_{jft}^{c} \tag{3.15}$$

where $\tilde{c}_{jt}^{f(r)} = c_{jt}^{f(r)}/USD\_SEK_t$ correspond to the estimated marginal costs, expressed in US dollars per DDD (USD_SEK_t being the exchange rate), λ_j is a product fixed effect and $Time_t$ a quarterly time trend. ε_{jft}^{c} represents an idiosyncratic cost shock. In this equation, the key explanatory variable driving learning-by-doing is the experience in production, e_{jft}, which needs to be computed. From section 3.4.2, it is specified as:

$$e_{jft} = \sum_{f' \in \mathcal{F}_m} w_{f'} \left(\sum_{\tau=0}^{t-1} \sum_{r \in R_{j\tau}, f(r)=f'} s_{jr\tau}(p_{jr\tau}) M_{r\tau} \right) \tag{3.16}$$

where $s_{jrt}(p_{jrt})M_{rt}$ is recovered from sales data. $\tau = 0$ corresponds to market entry, such that we rely on sales data from 1997 to compute product-level experience. We assume $s_{jrt}(p_{jrt})M_{rt}$ is observed for all $r \in R_{jt}$. The production facility for units sold in country r, $f(r)$, is imputed from the geographical distance between registered manufacturing sites, $f \in \mathcal{F}_m$, and the country's main city. The relative weights of experience at the same or different factories, w_f, are unobserved. We estimate the marginal cost model for different values of w_f, including $w_f = 1$, $\forall f \in \mathcal{F}_m$ and $w_f = 1$ if $f = f(r)$, 0 otherwise. The production in period t, Q_{jft}, is given by

$$Q_{jft} = \sum_{r \in R_{jt}, f(r)=f} \left(s_{jrt}(p_{jrt}) M_{rt} \right) \tag{3.17}$$

Quantities depend on equilibrium prices such that the model given by equation (3.15) cannot be estimated by OLS. We rely on an instrumental variable approach to correct for the endogeneity of current production volumes. We use the population-weighted average number of

competitors for product j in quarter t and country r, competitors within the same molecule, and competitors sharing the same characteristics as demand-side shifters for Q_{jft}. We assume that the unobserved cost shocks, ε^c_{jft}, are uncorrelated over time.[19]

The results are presented in Table 3.4. The first column presents the OLS estimates, which abstract away from the endogeneity of Q_{jft}. Columns 2 to 6 account for the endogeneity of current volumes, Q_{jft}, using demand shifters as instruments. In column 3, we allow for heterogeneous returns to experience between branded and biosimilar drugs. In column 4, we test for age as an alternative driver of learning. In our baseline specifications (columns 1 to 4), experience and volumes are measured across factories, assuming all units are produced in the same factory. Columns 5 and 6 measure experience and production exclusively at the active substance manufacturer level (Column 5), batch release location (Column 6). Finally, column 7 tests that the results are not driven by the Swedish market, excluding Swedish volumes from experience and volumes. The comparison of the first two columns confirms the endogeneity of Q_{jft}. From column 2 onwards, the estimates for γ suggest that economies of scale are small, yet not significant. $\hat{\eta}$ is negative and statistically significant in all specifications, providing support for learning-by-doing. The comparison of columns (3) and (4) suggests that measuring experience as a function of past production volumes matters beyond the age of the product, as the coefficient for $log(age_{jt})$ is small and non-significant once we control for past production. Learning-by-going is not only driven by time passing while producing product j: the intensive margin matters. There is a slight difference in returns to experience for branded and biosimilars (Column 4), and the result remains stable throughout the experience definition (Columns 5, 6, and 7). Column 4 is our preferred specification. In this specification, when experience doubles, the marginal cost of production decreases by 6.4% for branded drugs, 10.3% for biosimilars, suggesting that learning-by-doing happens slowly.[20] Yet, it should be interpreted with respect to the unit that we consider: one daily dose of a drug. Figure 3.1 plots the product fixed effects, suggesting a bit of heterogeneity across products. The coefficient for branded products is never significantly lower than that for the biosimilar counterparts, suggesting that branded manufacturers are, everything else being equal, no more efficient than biosimilar manufacturers. This could be driven by the extent to which earlier entry of the branded drug happens with

[19] This assumption is at odds with the production function literature (Olley and Pakes (1992)). In our context, any correlation in the unobserved cost shocks over time leads to an endogeneity issue as experience depends on past quantities. In future iterations, we will relax this assumption by instrumenting for both experience and current production volumes.

[20] The progress ratio is $\rho = 2^{\eta}$. The marginal cost decreases by $(1 - \rho) * 100$ percent when the experience stock doubles.

Table 3.4: Cost function estimation

	(1) OLS	(2) IV	(3) IV	(4) IV	(5) IV Active S.	(6) IV Batch R.	(7) IV excl. Sweden
$\log(Q_{jft})(\gamma)$	0.269***	-0.0970	-0.0821	-0.0419			
	(0.0901)	(0.118)	(0.119)	(0.112)			
$\log(e_{jft})(\eta)$	-0.261***	-0.114**	-0.156*				
	(0.0458)	(0.0474)	(0.0874)				
$\log age$			0.0679				
			(0.123)				
$\log(e_{jft})(\eta)$ non biosimilar				-0.0962**			
				(0.0487)			
$\log(e_{jft})(\eta)$ biosimilar				-0.157***			
				(0.0491)			
$\log(Q^{AS}_{jft})(\gamma)$					-0.142*		
					(0.0811)		
$\log(e^{AS}_{jft})(\eta)$ non biosimilar					-0.0454		
					(0.0365)		
$\log(e^{AS}_{jft})(\eta)$ biosimilar					-0.117***		
					(0.0407)		
$\log(Q^{BR}_{jft})(\gamma)$						-0.0362	
						(0.109)	
$\log(e^{BR}_{jft})(\eta)$ non biosimilar						-0.0984**	
						(0.0476)	
$\log(e^{BR}_{jft})(\eta)$ biosimilar						-0.167***	
						(0.0502)	
$\log(Q^{Excl.Sweden}_{jft})(\gamma)$							-0.0734
							(0.116)
$\log(e^{Excl.Sweden}_{jft})(\eta)$ non biosimilar							-0.0910*
							(0.0497)
$\log(e^{Excl.Sweden}_{jft})(\eta)$ biosimilar							-0.145***
							(0.0501)
Product FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time	Trend	Trend	Trend	Trend	Trend	Trend	Trend
Constant	4.240***	7.619***	8.055***	6.282***	6.983***	6.228***	6.723***
	(0.776)	(1.213)	(1.521)	(1.115)	(0.896)	(1.086)	(1.162)
Observations	438	438	438	438	438	438	437
R-squared	0.775	0.764	0.765	0.769	0.764	0.771	0.768
Weak IV Fstat		34.26	37.09	42.90	127.4	42.56	40.86
Volume	All	All	All	All	Active Subst.	Batch R.	Excl. Sweden

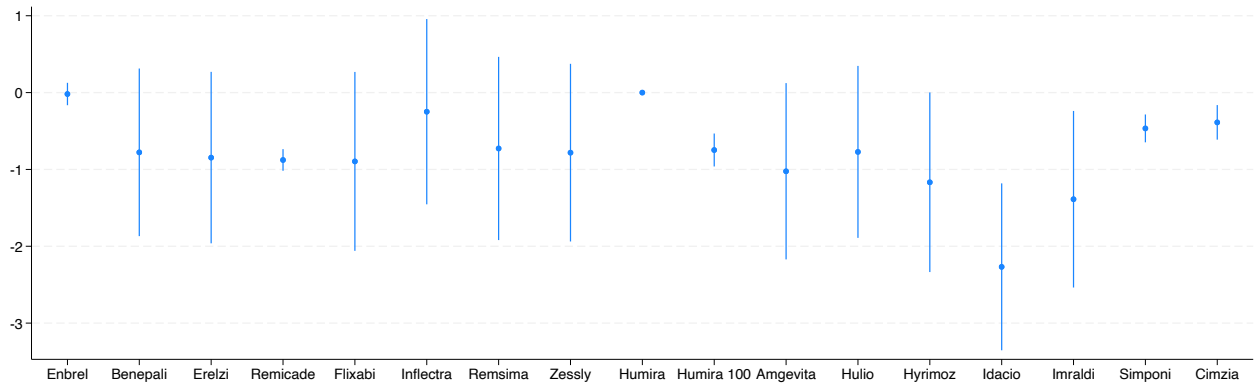
Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

facilities that are older and eventually outdated at the time of biosimilar entry, limiting the scope of cost efficiencies for the originator (Blackstone and Joseph (2013)).

3.6 Learning-by-doing and efficiency gains

In this section, the estimates from the cost function are used to assess the impact of learning-by-doing on product-level marginal cost. We perform three types of counterfactual scenarios, trying to understand (*i*) the heterogeneity in the contribution of countries to efficiency gains; (*ii*) the consequences for delaying biosimilar entry, and (*iii*) when the country delaying entry is large (eg, the US). To answer these questions, we use the estimates from Table 3.3, column

Figure 3.1: Cost function product fixed effect



4, and compute marginal cost levels prevailing under alternative experience levels, $\tilde{e}_{jft}$, and compare them to the estimated ones. In each scenario, we make the simplifying assumption that quantities in other countries and, where applicable, for other products would have remained constant. Moreover, as economies of scale are small and non-significant, we do not account for the potential impact on the production scale.

3.6.1 Country-level contribution

In this paragraph, we consider the contribution of individual countries to cost efficiencies in the production of biosimilars. We consider independently the contributions of France, Germany, the UK, and Sweden to cost efficiencies. In each scenario, denoted C_r, we consider the alternative level of experience that would have occurred had the product never entered country r. In counterfactual C_r, the level of experience for product j in factory f at time t becomes

$$\tilde{e}_{jft}^{C_r} = \sum_{f' \in \mathcal{F}_m} w_{f'} \left(\sum_{\tau=0}^{t-1} \sum_{r' \in R_{j\tau} \setminus \{r\},\ f(r')=f'} s_{jr'\tau}(p_{jr'\tau}) M_{r'\tau} \right) \tag{3.18}$$

Figure 3.2 plots the percentage change in marginal costs as products did not enter country r (on the horizontal axis), relative to the observed values at the end of the sample period. The effect ranges from nearly zero to +14.4%, suggesting asymmetries across products and countries. Had biosimilars not entered the German market, the marginal cost in the second quarter of 2022 would have been 4% higher on average, compared to 0.6% for Sweden. The effect is not constant over time. In the case of Germany, marginal costs would have been up to 10% higher soon after the drug's introduction (Figure 3.3). The last box of Figure 3.2 considers the impact of the biosimilar entering the US market. Due to patent disputes, over our sample period, biosimilars

only entered in one out of three molecules, Infliximab.[21] Given the considerable size of the US market, absent the US experience, the marginal cost for the biosimilar entering both the US and European market would have been 5.7% higher.

3.6.2 Delaying biosimilar entry

As marginal cost decreases with experience, delays in the entry of biosimilars affect efficiency gains across countries, leading to higher marginal costs and potentially higher prices even in countries subject to biosimilar competition. In Canada, all but one biosimilar entering the market entered after the product was launched in Europe. The average delay between the first launch and entry into Canada was 21 months, and the marginal cost was 64.5% lower when the product entered Canada compared to the EU. To measure the impact of delaying biosimilar entry on cost efficiencies, we calculate the marginal costs that would have prevailed had the product been launched in Canada simultaneously with its entry into the EU. In this simplified counterfactual, we assume that the adoption rate of biosimilars in Canada would have been the same under the alternative entry dates.[22] For each biosimilar, we consider the difference between the marginal cost under this alternative entry scheme and the observed one. We do not extrapolate the Canadian consumption volume beyond the observed quantities and consider, for each product, the last period in which Canadian volumes were observed. Figure 3.4 plots the corresponding distribution. Costs would have been up to 7.4% lower had the product entry in Canada not been delayed.

3.7 Conclusion

This paper quantifies learning-by-doing in the production process of biologic drugs. As manufacturing costs data for pharmaceuticals are unavailable, estimating the marginal costs of biologic drugs requires using a demand and pricing approach. To that end, we develop a tractable model of biologic drug supply and demand where, on the supply side, the production process embeds learning-by-doing and economies of scale. We estimate the model for TNF-α inhibitors, combining information on drug consumption across 25 countries and production facilities. As the pricing game is inherently complex to solve, we recover marginal costs from the pricing

[21] Only one biosimilar for Infliximab was available both in the US and Sweden over our sample period.
[22] This might not be the case in practice, given that the higher marginal cost level at the time of entry can lead to higher prices.

Figure 3.2: Country-level impact on efficiency gains

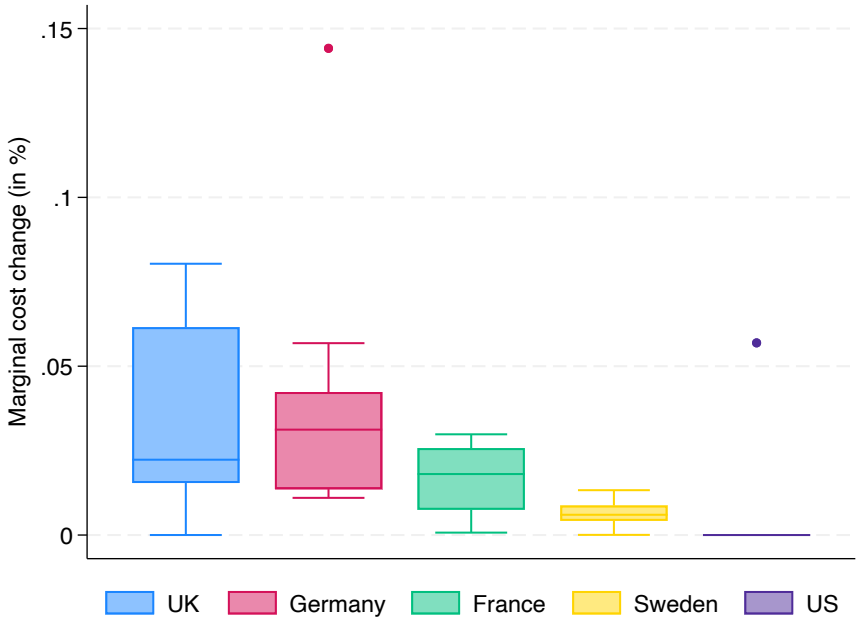
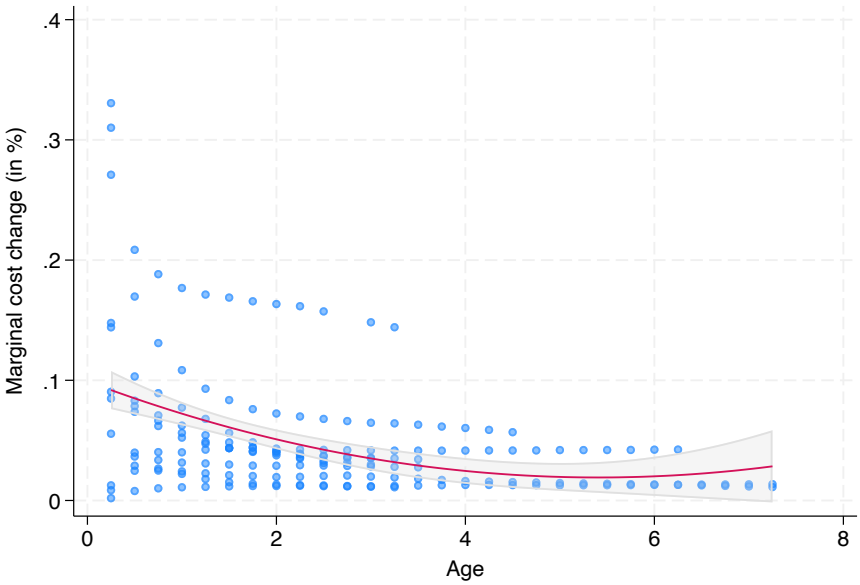
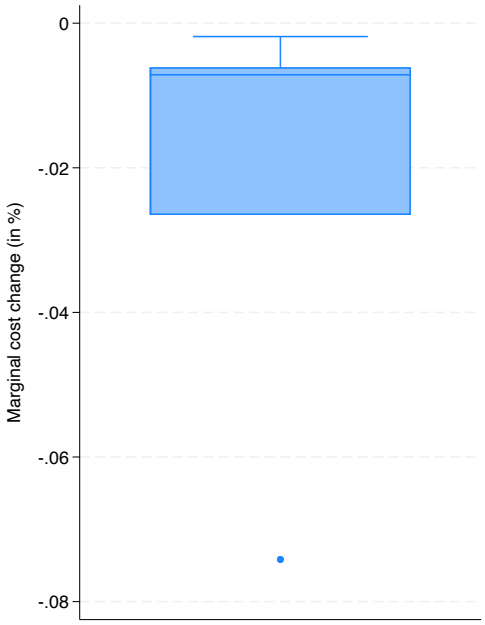


Figure 3.3: German volumes impact over the product age



Note:

Figure 3.4: Country-level impact on efficiency gains



game in the Swedish market. We assume that observed prices in Sweden are close to the static Bertrand-Nash prices, given the small contribution of Swedish volumes to learning-by-doing.

The structural estimation results suggest significant marginal costs decrease over the product's lifespan. On average, the marginal cost of TNF-α inhibitors' biosimilars has decreased by 46% since market entry. The marginal cost function estimates provide evidence for learning-by-doing, while economies of scale are small and not statistically significant. The results suggest that the marginal cost for biosimilars decreases by 10.3% every time experience doubles. The concentration of the production process implies that country-level biosimilar penetration matters for cost efficiencies in other countries. The size of this spillover depends on the size of the market. Without entry in the German market, biosimilar costs would have been 4% higher at the end of the sample period, compared to 0.6% without entry in the Swedish market. Learning-by-doing also implies that delaying biosimilar entry in a given market impacts efficiency gains in other markets. Biosimilar entry in Canada occurred on average 21 months after the entry in the EU market, when marginal costs were, on average, 64.5% lower than at the time of EU entry. Had biosimilar entry been simultaneous, marginal costs would have been on average 2.4% lower 15 months after market entry.

Bibliography

Arrow, K. J. (1962). The economic implications of learning by doing. *The review of economic studies*, 29(3):155–173.

Benkard, C. L. (2000). Learning and forgetting: The dynamics of aircraft production. *American Economic Review*, 90(4):1034–1054.

Benkard, C. L. (2004). A dynamic analysis of the market for wide-bodied commercial aircraft. *The Review of Economic Studies*, 71(3):581–611.

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

Besanko, D., Doraszelski, U., Kryukov, Y., and Satterthwaite, M. (2010). Learning-by-doing, organizational forgetting, and industry dynamics. *Econometrica*, 78(2):453–508.

Blackstone, E. A. and Joseph, P. F. (2013). The economics of biosimilars. *American health & drug benefits*, 6(8):469.

Cabral, L. M. and Riordan, M. H. (1994). The learning curve, market dominance, and predatory pricing. *Econometrica: Journal of the Econometric Society*, pages 1115–1140.

Coghlan, J., He, H., and Schwendeman, A. S. (2021). Overview of humira® biosimilars: current european landscape and future implications. *Journal of pharmaceutical sciences*, 110(4):1572–1582.

De Loecker, J. and Scott, P. (2024). Markup estimation using production and demand data: An application to the us brewing industry. *accepted at the Review of Economic Studies*.

Doraszelski, U. and Pakes, A. (2007). A framework for applied dynamic analysis in io. *Handbook of industrial organization*, 3:1887–1966.

Feng, J., Hwang, T., Klimek, J., and Maini, L. (2024). Biosimilar entry and the pricing of biologic drugs.

Fudenberg, D. and Tirole, J. (1983). Learning-by-doing and market performance. *The Bell Journal of Economics*, pages 522–530.

Khmelnitskaya, E., Marshall, G., and Orr, S. (2024). Identifying scale and scope economies using product market data. *Available at SSRN 4757085*.

Levitt, S. D., List, J. A., and Syverson, C. (2013). Toward an understanding of learning by doing: Evidence from an automobile assembly plant. *Journal of political Economy*, 121(4):643–681.

MacKay, A. and Remer, M. (2024). Consumer inertia and market power. *Available at SSRN 3380390*.

Maini, L. and Pammolli, F. (2023). Reference pricing as a deterrent to entry: evidence from the european pharmaceutical market. *American Economic Journal: Microeconomics*, 15(2):345–383.

Olley, S. and Pakes, A. (1992). The dynamics of productivity in the telecommunications equipment industry.

Pontén, J., Rönnholm, G., Skiöld, P., et al. (2017). Ppri pharma profile sweden 2017. *Stockholm: the Dental and Pharmaceutical Benefits Agency*.

Rémuzat, C., Urbinati, D., Mzoughi, O., El Hammi, E., Belgaied, W., and Toumi, M. (2015). Overview of external reference pricing systems in europe. *Journal of market access & health policy*, 3(1):27675.

Scott Morton, F. M., Stern, A. D., and Stern, S. (2018). The impact of the entry of biosimilars: evidence from europe. *Review of Industrial Organization*, 53:173–210.

Thompson, P. (2010). Learning by doing. *Handbook of the Economics of Innovation*, 1:429–476.

Thompson, P. (2012). The relationship between unit cost and cumulative quantity and the evidence for organizational learning-by-doing. *Journal of Economic perspectives*, 26(3):203–224.

3.8 Additional tables and figures

Table 3.5: Country-level summary statistics, Fixed set (IQVIA data)

Country	In 2022	In 2013	Molecules	Products	Biosimilars	Per molecule		Country share		Within country
						Product	Biosimilar	Value	Volume	Biosimilar share
				(#)	(#)	(#)	(#)	(%)	(%)	(%)
US	1	1	13	54	21	3	1	70	45	6.8
Germany	1	1	13	74	51	6	4	5.9	8.2	14
France	1	1	13	63	44	5	3	4.3	6.1	12
Canada	1	1	12	52	26	4	2	3.3	2.9	8.8
Italy	1	1	13	57	41	4	3	3.0	6.3	44
UK	1	1	13	61	36	4	2	3.0	5.2	28
Spain	1	1	13	57	40	4	3	2.9	5.6	17
Belgium	1	1	13	46	32	4	3	0.9	0.8	4.2
Sweden	1	1	13	53	35	4	3	0.8	0.9	19
Russia	1	1	13	53	7	3	0	0.7	3.7	0.9
Switzerland	1	1	13	51	32	4	2	0.7	0.4	4.2
Turkey	1	1	13	38	13	3	1	0.6	7.2	1.8
Austria	1	1	13	59	43	4	3	0.5	0.6	49
Ireland	1	1	12	45	32	3.5	2.5	0.5	0.3	8.5
Finland	1	1	13	52	36	4	3	0.4	0.9	12
Norway	1	1	12	37	23	3	2	0.4	0.3	21
Poland	1	1	13	47	34	3	3	0.4	2.6	9.7
Czech Republic	1	1	13	54	39	4	2	0.3	0.5	15
Hungary	1	1	12	42	28	3.5	2	0.3	0.4	13
Portugal	1	1	13	53	36	4	3	0.3	0.7	21
Greece	1	1	11	29	16	2	1	0.2	0.8	8.6
Bulgaria	1	1	13	43	28	3	2	0.1	0.5	13
Croatia	1	1	13	55	41	4	3	0.09	0.1	32
Slovenia	1	1	12	43	30	3	2	0.07	0.08	13
Luxembourg	1	1	10	22	13	2	1	0.02	0.03	0.5
All	.	.	13	50	31	4	2	100	100	12

Note: Column 3 corresponds to the number of molecules available in country c. Column 4 and 5 contain the number of products and biosimilars. Columns 6 (resp. 7) the median number of products (resp. biosimilars) per molecule. Columns 8 and 9 show the contribution of country c to the total value and volume. Column 10 corresponds to the share of biosimilar units within a country.

Figure 3.5: Biologic drugs approval over time

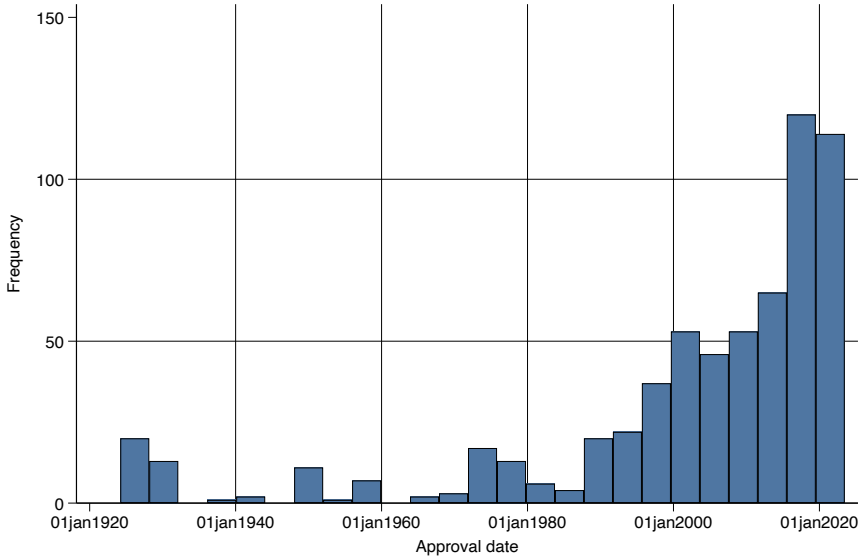
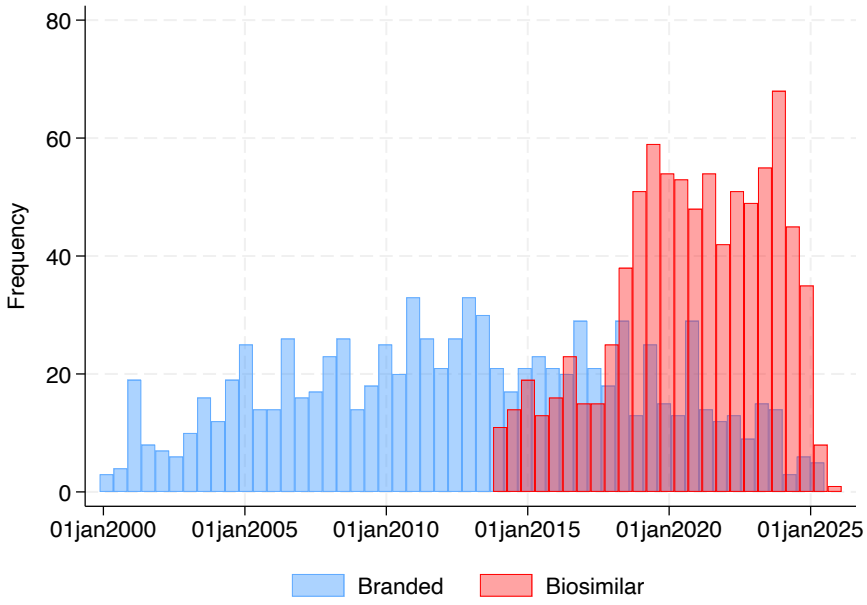


Figure 3.6: Changes implemented after the marketing authorization



Note: Number of distinct applications submited to the EMA over time.

Table 3.6: Country-level summary statistics, All countries (IQVIA data)

Country	In 2022	In 2013	Molecules	Products	Biosimilars	Per molecule		Country share		Within country
						Product	Biosimilar	Value	Volume	Biosimilar share
				(#)	(#)	(#)	(#)	(%)	(%)	(%)
US	1	1	13	54	21	3	1	68	42	6.8
Germany	1	1	13	74	51	6	4	5.7	7.8	14
France	1	1	13	63	44	5	3	4.2	5.8	12
Canada	1	1	12	52	26	4	2	3.2	2.8	8.8
Italy	1	1	13	57	41	4	3	2.9	6.0	44
UK	1	1	13	61	36	4	2	2.9	5.0	28
Spain	1	1	13	57	40	4	3	2.8	5.4	17
Netherlands	1	0	13	51	34	4	3	1.0	1.1	16
Belgium	1	1	13	46	32	4	3	0.8	0.7	4.2
Japan	0	1	9	9	0	1	0	0.8	0.5	0
Sweden	1	1	13	53	35	4	3	0.7	0.9	19
Russia	1	1	13	53	7	3	0	0.7	3.5	0.9
Switzerland	1	1	13	51	32	4	2	0.6	0.4	4.2
Turkey	1	1	13	38	13	3	1	0.6	6.9	1.8
Austria	1	1	13	59	43	4	3	0.5	0.6	49
Ireland	1	1	12	45	32	3.5	2.5	0.4	0.3	8.5
Finland	1	1	13	52	36	4	3	0.4	0.8	12
Norway	1	1	12	37	23	3	2	0.4	0.3	21
Poland	1	1	13	47	34	3	3	0.4	2.5	9.7
Denmark	1	0	13	43	28	4	2	0.4	0.3	26
Czech Republic	1	1	13	54	39	4	2	0.3	0.5	15
Hungary	1	1	12	42	28	3.5	2	0.3	0.4	13
Romania	1	0	12	45	31	3.5	2.5	0.3	0.9	9.0
Portugal	1	1	13	53	36	4	3	0.3	0.6	21
Australia	0	1	12	15	1	1	0	0.3	0.3	0
Greece	1	1	11	29	16	2	1	0.2	0.7	8.6
Slovakia	1	0	12	52	39	4	3	0.2	0.2	29
Bulgaria	1	1	13	43	28	3	2	0.1	0.5	13
Croatia	1	1	13	55	41	4	3	0.09	0.1	32
Slovenia	1	1	12	43	30	3	2	0.07	0.08	13
Serbia	1	0	13	32	19	3	2	0.04	0.2	7.1
Korea	0	1	9	9	0	1	0	0.04	0.1	0
China	0	1	8	8	0	1	0	0.04	0.09	0
Ukraine	1	0	12	48	13	2.5	1	0.03	0.6	3.2
Thailand	0	1	10	10	0	1	0	0.03	0.06	0
Bosnia	1	0	13	28	16	2	1	0.02	0.08	16
South Africa	0	1	9	9	0	1	0	0.02	0.06	0
Luxembourg	1	1	10	22	13	2	1	0.02	0.03	0.5
Brazil	0	1	10	10	0	1	0	0.02	0.07	0
Belarus	1	0	12	30	2	2.5	0	0.02	0.1	1.4
Mexico	0	1	10	11	1	1	0	0.02	0.04	0.07
New Zealand	0	1	9	9	0	1	0	0.01	0.01	0
Singapore	0	1	9	9	0	1	0	0.01	0.007	0
India	0	1	8	9	1	1	0	0.009	0.08	0
Algeria	0	1	1	1	0	1	0	0.009	0.08	0
Argentina	0	1	11	11	1	1	0	0.007	0.05	0
Malaysia	0	1	9	9	0	1	0	0.006	0.01	0
Venezuela	0	1	5	5	0	1	0	0.005	0.01	0
Philippines	0	1	9	9	0	1	0	0.005	0.01	0
UAE	0	1	9	9	0	1	0	0.004	0.009	0
Saudi Arabia	0	1	8	8	0	1	0	0.004	0.02	0
Indonesia	0	1	7	7	0	1	0	0.004	0.01	0
Latvia	0	1	10	11	2	1	0	0.002	0.009	0
Pakistan	0	1	8	8	0	1	0	0.002	0.006	0
Colombia	0	1	10	10	0	1	0	0.002	0.005	0
Ecuador	0	1	10	10	0	1	0	0.002	0.02	0
Egypt	0	1	5	5	0	1	0	0.002	0.03	0
Lebanon	0	1	9	9	1	1	0	0.002	0.008	0.2
Uruguay	0	1	9	9	0	1	0	0.001	0.008	0
Tunisia	0	1	1	1	0	1	0	0.001	0.01	0
Peru	0	1	5	5	0	1	0	0.0007	0.002	0
Morocco	0	1	1	1	0	1	0	0.0006	0.006	0
Kuwait	0	1	7	7	0	1	0	0.0002	0.0009	0
Jordan	0	1	3	3	0	1	0	0.0002	0.0010	0
All	.	.	10	28	15	2	0	100	100	12

Note: Column 3 corresponds to the number of molecules available in country c. Columns 4 and 5 contain the number of products and biosimilars. Columns 6 (resp. 7) the median number of products (resp. biosimilars) per molecule. Columns 8 and 9 show the contribution of country c to the total value and volume. Column 10 corresponds to the share of biosimilar units within a country.

Table 3.7: Molecule level summary statistics

ATC 4	Molecule	**1** Chronic	Volume (%)	Value (%)	Hospital (%)	Biosimilars Entry	Biosimilars Products (#)	Biosimilars Volume (%)	Prices Branded Av.	Prices Biosimilar Av.
A10C5	Insulin Glargine	1	53	13	9.0	2015	2	7.7	42	38
B1B2	Enoxaparin Sodium	0	27	1.2	53	2017	1	13	3.4	3.4
B3C0	Epoetin Alfa	1	2.3	0.7	58	2007	4	72	15	68
H4C0	Somatropin	1	0.3	0.5	32	2006	1	68	239	254
H4E0	Teriparatide	1	0.3	1.6	6.6	2019	3	2.7	1135	264
L1G1	Rituximab	1	1.2	9.4	61	2017	5	12	1458	879
L1G2	Bevacizumab	1	1.0	7.5	68	2019	6	9.3	864	950
L1G3	Trastuzumab	1	1.0	7.4	66	2018	6	9.2	1126	1050
L3A1	Filgrastim	0	1.4	0.7	60	2006	5	72	34	85
L3A1	Pegfilgrastim	0	0.5	7.6	52	2018	8	8.3	1998	1288
L4B0	Adalimumab	1	3.4	25	22	2018	8	9.0	1761	343
L4B0	Etanercept	1	5.6	14	15	2016	3	6.7	743	181
L4B0	Infliximab	1	3.1	12	60	2013	5	14	440	385

Note: Volume (resp. Value) corresponds to the contribution of the molecule to total volume (resp. value). Hospital corresponds to the share sold to the hospital sector (vs retail). Biosimilars entry, products and volume the the year of the firt biosimilar entry, the number of biosimilars and their volume within the molecule. The last two colums correspond to the average price of branded and biosimilar products.

Table 3.8: Approved facilities, TNF-α inhibitors

Molecule	Biosimilar	EMA Active substance Country	EMA Active substance Locations	EMA Batch release Country	EMA Batch release Locations	**1** FDA	FDA Active substance Confidential	FDA Active substance locations	FDA Batch release Confidential	FDA Batch release Locations
Adalimumab	0	3	4	1	1	1	1	.	1	.
	1	1.12	1.25	1.88	2.25	0.62	0.20	1	0.80	1
Certolizumab Pegol	0	1	1	1	1	1	1	.	1	.
Etanercept	0	1	1	1	1	1	1	.	1	.
	1	1.67	1.67	1.33	1.67	0.33	0	1	0	1
Golimumab	0	2	2	1	1	1	1	.	1	.
Infliximab	0	2	2	1	1	1	1	.	1	.
	1	1.50	1.75	1.75	2.25	0.25	0	1	1	.

Note: This table summarizes production facilities authorized by the EMA and FDA (if available). It breaks down information about active substance manufacturing locations and batch release locations. Columns 2 and 4 provide the average number of facilities authorized by the EMA. Columns 5-9 summarize the extent of information extracted from FDA sources. Importantly, not all biosimilars are available in the US (column 5), of those available, information about active substance and batch release is often missing (columns 6 and 8).

Table 3.9: Product launches across countries and sales data of TNF-α inhibitors

Molecule	Product	Biosimilar	Countries		Population	
			Number (#)	In sales data (Share)	Size (M hab.)	In sales data (Share)
Etanercept	Enbrel	0	55	0.51	5,244	0.20
Adalimumab	Humira	0	63	0.49	3,696	0.32
Infliximab	Inflectra	1	54	0.54	3,543	0.32
Etanercept	Benepali	1	47	0.62	3,145	0.34
Infliximab	Remicade	0	47	0.57	3,116	0.29
Adalimumab	Amgevita	1	40	0.68	2,734	0.34
Etanercept	Erelzi	1	35	0.77	2,480	0.42
Etanercept	Nepexto	1	35	0.71	2,152	0.26
Adalimumab	Cyltezo	1	45	0.62	1,787	0.58
Infliximab	Flixabi	1	40	0.72	1,488	0.78
Adalimumab	Hukyndra	1	43	0.65	1,324	0.74
Adalimumab	Hulio	1	40	0.72	1,291	0.83
Adalimumab	Idacio	1	40	0.68	1,232	0.84
Adalimumab	Hyrimoz	1	37	0.73	1,204	0.86
Adalimumab	Imraldi	1	36	0.75	1,120	0.92
Adalimumab	Yuflyma	1	36	0.72	1,101	0.81
Infliximab	Zessly	1	35	0.77	1,060	0.87
Adalimumab	Amsparity	1	34	0.76	925	0.96

Note: This table compares the countries in which products are launched and the countries available in our most recent sales dataset (2010-2022). Products are ordered by the size of the population in countries in which it is launched. Column 1 corresponds to the number of countries in which the product is launched. Column 2 contains the share of countries in which the product is launched, which are included in the sales data sample. Column 3 corresponds to the total population in regions where the product is available, column 4 to the share of the population living in countries where the drug is approved and are in the sales data. Drug entries are observed with the Citeline Pharmaproject data.

Table 3.10: Determinants of prices for TNF-α inhibitors with regulated prices

	(1) France	(2) Germany	(3) UK
α_s non biosimilar	0.0609**	0.0158	0.00576**
	(0.0266)	(0.0200)	(0.00228)
α_s biosimilar	-0.0106	0.0111	0.00528
	(0.0108)	(0.00905)	(0.00437)
Log(age) (γ)	-0.0646	-0.0676*	-0.0238
	(0.0470)	(0.0398)	(0.0216)
Nb products (β_1)	-0.0175***	-0.0599***	-0.00417
	(0.00591)	(0.00862)	(0.00255)
Exchange rate	-1.752***	-1.653***	-1.653***
	(0.139)	(0.126)	(0.0851)
Time FE	Trend	Trend	Trend
Year x Country FE	No	No	No
Molecule FE	No	No	No
Product FE	Yes	Yes	Yes
Biosimilar x Country FE	No	No	No
Molecule x Year FE	Trend	Trend	Trend
Constant	78.85***	-11.63	-15.01**
	(14.73)	(12.99)	(5.877)
Observations	467	491	480
R-squared	0.945	0.904	0.981

Note: Robust standard errors in parenthesis.

Table 3.11: Molecules in TNF-α inhibitors, L04AB

ATC5	Molecule	Biosim. Entry	Injection	Frequency	Crohn
L04AB01	Etanercept	10/2016	Self	Once/Twice a week	
L04AB02	Infliximab	1/2015	Hosp. until 2021	Every 4/8 weeks	✓
L04AB04	Adalimumab	11/2018	Self	Every 2 weeks	✓
L04AB05	Certolizumab pegol			Every month	✓
L04AB06	Golimumab			Every month	

Table 3.12: Adalimumab product characteristics

Product	Laboratory	Entry date	Citrate-free	Concentration	Needle size	Latex-free
Humira	Abbvie			50	Big	
Humira v2	Abbvie	2016	✓	100	Small	✓
Amgevita	Amgen	2018	✓	50	Small/Big	
Imraldi	Biogen	2018	(✓from 11-2022)	50	Small	✓
Hulio	Viatris	2019	✓	50	Small	✓
Hyrimoz	Sandoz	2019		50	Big	
Hyrimoz v2[a]	Sandoz	2023	✓	100	Big	
Idacio	Fresenius	2019		50	Small	✓
Yuflyma*	Celltrion	2021	✓	100	Small	✓
Amsparity	Pfizer	2021	✓	50	N/A	✓
Hukyndra*	EG	2022	✓	100	Small	✓

Source: Coghlan et al. (2021). * Information from other sources. [a]New version of Hyrimoz end of 2023: critrate free and high concentration. https://www.sandoz.com/sandoz-launches-hyrimozr-adalimumab-high-concentration-formulation-europe-aiming-improve-patient/

Table 3.13: EMA post-authorization variations by type, TNF-α inhibitors

Type	Output	Element	Share (%)
Product incl.			0.466
	Active Substance	Container	0.008
		Control	0.182
		Design	0.013
		Manufacturing	0.161
		Manufacturing Sites	0.051
	Final Product	Composition	0.004
		Container	0.116
		Control	0.146
		Design	0.014
		Manufacturing	0.150
		Manufacturing Sites	0.153
		Missing	0.001
Label			0.167
Administrative			0.062
Other			0.095
Missing			0.211

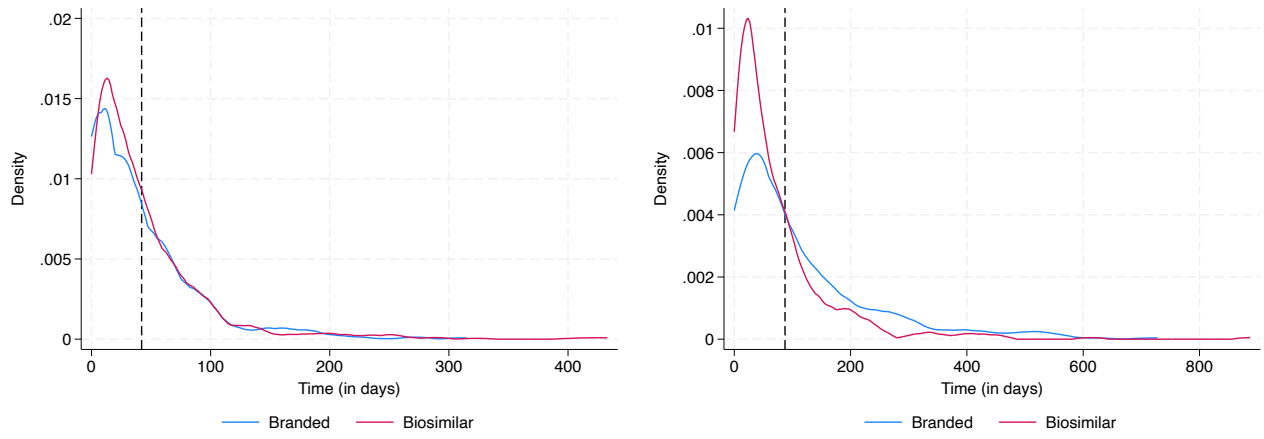
Note: This table presents the breakdown of variations implemented for TNF-α inhibitors from the product's marketing approval until 2025. 'Admin', 'Product', and 'Label' classes are based on paragraphs A, B.I, B.II, and C from the 2013 Guidelines. 'Product' gathers variations related to Active Substance (AS) and Final Product (FP) variations. 'Label' variations relate to the labelling of the product, 'Admin' corresponds to administrative changes. 'Other' includes variations related to paragraphs B.III, B.IV, B.V from these guidelines, marketing authorization transfer and renewal, unspecified quality changes, and labelling. 'Missing' represents variations that are not matched with one of these categories. Most of the changes in the 'Missing' categories were implemented prior to the guidelines.

Table 3.14: EMA post-authorization variations by molecule, TNF-α inhibitors

		A. Applications			B. Product variations			
Molecule	Biosimilar	Per quarter	Procedure/application	Total	Per quarter	Per product		
		Av.	Av.	(#)	Av.	Av.	Min	Max
Adalimumab	0	2.38	1.78	214	1.38	124	124	124
	1	1.48	1.62	39	1.72	43.6	2	76
Certolizumab Pegol	0	1.60	2.24	104	1.82	118	118	118
Etanercept	0	2.39	1.93	246	1.36	140	140	140
	1	1.71	1.54	57.3	1.54	50	38	61
Golimumab	0	1.74	1.98	113	1.34	87	87	87
Infliximab	0	2.31	1.51	240	1.07	111	111	111
	1	1.61	1.54	75	1.65	79.8	0	188

Note: Columns 1, 2, and 3 focus on the number of distinct applications, identified by their application number. Column 1 presents the average number of applications per product per quarter. Columns 4, 5, 6 and 7 focus on 'Product' variations (Table 3.13).

Figure 3.7: Time between two changes, in days

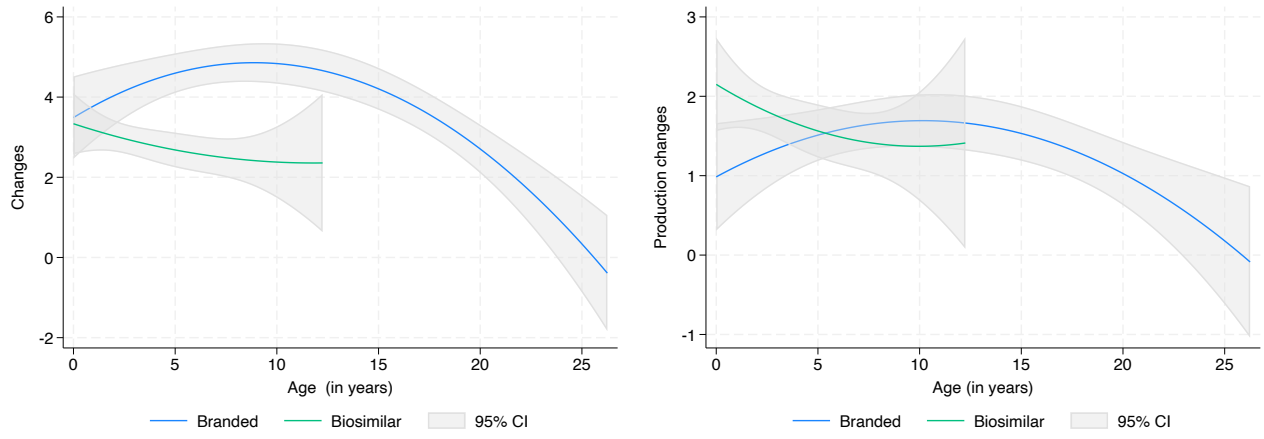


(a) All

(b) 'Product' variations

Note: These figures present the distribution of the time between two applications for the same product, separately for branded drugs and biosimilars. Figure 3.7a focuses on any change. Figure 3.7b focuses on 'Product' variation. The vertical dashed lines correspond to the average across both drug classes.

Figure 3.8: Quarterly variation through drug's age



(a) Any procedure

(b) 'Product' variations

Note: These figures present the evolution of variations over time since marketing approval, separately for branded drugs and biosimilars. Figure 3.8a focuses on any application. Figure 3.8b focuses on 'Product' variations.