Paying More at Fixed Prices: Pharmacy Incentives and Drug Substitution *

Haruo Kakehi

University of Wisconsin-Madison

Yoko Ibuka†

Keio University

August 26, 2026

Abstract

When regulated prices cannot clear shortages, non-price mechanisms determine who bears the cost. We study this in Japan's drug market, where pharmacists choose among differently priced substitutes, exploiting a government-ordered suspension of a major generic manufacturer. Out-of-pocket spending rose by up to 21%, driven by temporary shifts to brand-name drugs and larger, persistent shifts to higher-priced generics. Pharmacists' dispensing choices vary with financial incentives, yet the government's pay-for-performance program rewards generic use rather than lower-priced generics and cannot prevent within-generic cost increases. We evaluate targeted incentives to reduce spending, and find evidence of lower adherence and higher discontinuation of treatment.

Keywords: Supply disruptions; Drug prices; Generic drugs; Pharmacies; Drug shortages.
JEL codes: I11; I18; D12.

*This study was financially supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI grant-in-aid for scientific research B (24K00267). We thank participants at seminars at Kwansei Gakuin University and Kobe University for their valuable comments and suggestions. We are especially grateful to Naoki Wakamori for his thoughtful comments, which have substantially improved this paper.

†Corresponding Author.

1 Introduction

Supply disruptions are pervasive, and who bears their costs depends critically on whether prices can adjust. In flexible-price markets, scarcity is largely resolved through price increases; in regulated markets, it must instead be absorbed through non-price mechanisms such as rationing, queuing, or substitution across products. Such regulation is common in housing through rent control, in agriculture through price floors and ceilings, and in energy and health care through administratively set prices. Existing work studies how disruptions affect production, trade, and product availability (Barrot and Sauvagnat, 2016; Boehm et al., 2019; Carvalho et al., 2021), but less is known about how disruptions are transmitted to consumers when prices are fixed and product choice is mediated by an intermediary. We focus on substitution across close but differently priced products: when a disruption shifts consumers toward a more expensive one, their expenditure rises with no change in any product's price and an unintended cost that depends on which substitute the intermediary dispenses.

In health care markets, prices are often regulated and treatment choices are delegated to providers, whose incentives shape which products patients receive (Iizuka, 2012; Clemens and Gottlieb, 2014). Japan's prescription-drug market combines two features that are common in regulated markets but rarely observed together so clearly. First, products with the same active ingredient, dosage form, and strength are sold at different regulated prices. Second, pharmacists rather than patients often select the dispensed product[1]. Their financial incentives can shape whether a patient receives a brand-name product or a generic and, among generics, the price level of the generic dispensed. In line with generic-substitution policies worldwide, Japan's pay-for-performance program provides pharmacies with an add-on payment based on their generic-dispensing share, en-

[1]Patients may express a preference between brand-name and generic drugs but generally do not choose the manufacturer or price tier of the generic they receive. In Japan, advertising prescription drugs to the public is prohibited. Hospitals are independent of pharmacies, and physicians are prohibited from referring patients to specific pharmacies in any form.

couraging substitution toward generics rather than brand-name products to reduce drug spending. Because this add-on does not vary across generics, however, pharmacies have no incentive to select a lower-priced one. Intermediary incentives can thus shape not only whether substitution occurs but which substitute is dispensed, and hence how much of a disruption's cost ultimately falls on patients.

Drug shortages make these institutional features an immediate policy concern. Shortage risk is disproportionately concentrated in low-price, low-margin off-patent products, and the burden is large: recent estimates put the annual cost to EU community pharmacists at EUR 399–793 million (European Commission, 2025). This concentration creates a tension at the center of generic-drug policy. Generic entry and substitution lower patient and insurer spending (Iizuka, 2012; Ito et al., 2020; Janssen and Granlund, 2023), but shortages can shift patients from inexpensive generics to branded or higher-priced alternatives, reversing part of those savings. Existing work mainly studies why shortages arise and how manufacturers, providers, and prices respond (Ridley et al., 2016; Yurukoglu et al., 2017; Stomberg, 2018; Dubois et al., 2023; Conti and Wosińska, 2025); little causal evidence identifies which substitutes are ultimately dispensed or how patients' out-of-pocket costs change.

We provide this evidence by exploiting the government-ordered 2021 production suspension of Nichi-Iko Pharmaceutical, a major Japanese generic manufacturer, as a quasi-natural experiment. The suspension followed audits that uncovered manufacturing misconduct and was abrupt, unanticipated, and unrelated to patient demand or local market conditions. Using individual-level claims that link patients to their dispensing pharmacies, we compare patients who predominantly used Nichi-Iko products before the suspension with patients using generics from other manufacturers within the same drug group that shares active ingredient, dosage form, and strength (dose). A shortage creates two adjustment channels: patients may reduce or discontinue treatment, and, conditional on continued treatment, the pharmacy decides which product to dispense. We estimate how

the resulting shortage affects treatment continuity through the first channel and out-of-pocket spending and drug substitution through the second, among patients receiving treatment for dyslipidemia and hypertension. To interpret the substitution patterns, we model pharmacies as choosing among alternative products based on procurement costs, dispensing margins, and payments under the generic-dispensing pay-for-performance program, and examine whether responses differ by pharmacies' add-on status. Guided by the model, we further simulate a targeted within-generic incentive and show that a modest per-prescription payment can lower both patients' out-of-pocket spending and total medical spending, net of its cost.

We find that the supply shock increased patients' out-of-pocket spending by up to 21%. Event-study estimates reveal two substitution responses: a temporary shift from generics to brand-name drugs and a persistent shift toward higher-priced generics, with the latter accounting for most of the implied increase in patient spending. Its larger contribution reflects the much larger substitution response: far more prescriptions shift toward higher-priced generics than toward brand-name drugs. Its persistence is consistent with the generic-dispensing add-on dropping out of the choice among generics. Cross-pharmacy heterogeneity in the brand-versus-generic response is consistent with the program's incentive to remain within the generic segment. Our back-of-the-envelope calculations suggest that an additional financial incentive for dispensing lower-priced generics would more than pay for itself: an incentive of approximately JPY 60 per shortage-induced high-priced-generic prescription would reduce patient out-of-pocket spending by JPY 197 and total medical spending by JPY 663, net of the incentive cost. We also find class-specific evidence of lower adherence and higher discontinuation, indicating that the effects of shortages extend beyond financial costs to treatment continuity.

This paper makes three contributions. First, we contribute to the literature on the welfare consequences of price regulation and non-price adjustment. A large body of work has examined how price controls affect welfare theoretically (Bulow and Klemperer, 2012)

and empirically in markets such as housing (Glaeser and Luttmer, 2003) and energy (Frech and Lee, 1987). In health care, related work studies waiting time, priority rules, and reassignment as non-price mechanisms for allocating scarce resources (Gravelle and Siciliani, 2008; Agarwal et al., 2021; Huitfeldt et al., forthcoming). A common theme in this literature is that when prices cannot adjust to clear markets, supply disruptions are resolved through alternative allocation mechanisms rather than price increases. We contribute to this literature by studying pharmaceutical markets, where shortages are resolved through delegated substitution between clinically equivalent drugs with different prices. Unlike standard consumer markets, these substitution decisions are often made by pharmacists rather than patients themselves, implying that the incidence of shortages may depend not only on supply conditions but also on intermediary incentives. Second, we add to the limited economics literature on the consequences of drug shortages for patients. Prior studies have primarily examined the causes of shortages (Ridley et al., 2016; Stomberg, 2018; Dubois et al., 2023; Conti and Wosińska, 2025), with relatively little attention to their economic consequences. Because Japan's regulated setting shuts down the price channel, our estimates isolate the substitution channel alone, allowing us to identify the mechanism of cost increases more cleanly than in less regulated markets such as the United States, where price increases and substitution operate simultaneously. We further decompose this substitution into temporary shifts toward brand-name drugs and persistent shifts toward higher-priced generics, and document additional effects on treatment continuity. Third, we contribute to the understanding of provider behavior in healthcare markets by examining how pharmacy dispensing responses vary under a pay-for-performance program. While extensive research has analyzed physician responses to financial incentives (Chalkley and Tilley, 2005; Iizuka, 2012; Chan et al., 2022), the role of pharmacies remains relatively understudied despite growing attention to generic drugs and pharmacy markets (Brekke et al., 2013; Ito et al., 2020; Janssen and Granlund, 2023; Kakehi and Nakajima, 2026). We distinguish the brand-versus-generic margin from the

choice among generics: the existing program rewards generic dispensing but provides no direct incentive to select a lower-priced generic. We therefore evaluate a targeted within-generic payment and quantify its implications for patient out-of-pocket spending and total drug spending.

2 Background

2.1 Production Suspension and Drug Shortages

To define the supply shock, our analysis centers on the government mandate issued in early 2021 requiring the generic drug manufacturer Nichi-Iko Pharmaceutical Co., Ltd. (hereafter NP) to suspend production in response to manufacturing misconduct (Nichi-Iko Co., Ltd., 2021).

We focus on the production suspension of the manufacturer for two reasons. First, it was one of the first large-scale suspensions in the wave of administrative sanctions from 2021–2023 and was therefore especially unexpected relative to later actions. Indeed, these large-scale scandals became major events that heightened concerns about the stability of supply and the quality of generic drugs. Second, NP is a leading generic-drug manufacturer in Japan, supplying a wide range of therapeutic categories and product variants. In fiscal year 2020, it reported consolidated revenue of JPY 188.2 billion and marketed over 1,000 pharmaceutical products, giving it one of the broadest product portfolios among Japanese generic manufacturers (Nichi-Iko Pharmaceutical Co., Ltd., 2021; Nichi-Iko Pharmaceutical Co., Ltd. and Eisai Co., Ltd., 2019). The firm's importance became evident following the suspension of its manufacturing operations in 2021 [2]. The

[2]The immediate regulatory context for NP's suspension was set by a high-profile incident at another generic manufacturer. In December 2020, Kobayashi Kako Co., Ltd. was found to have distributed antifungal tablets contaminated with a sedative agent, resulting in serious health injuries including traffic accidents; the company subsequently received a record 116-day production suspension order in February 2021. Heightened regulatory scrutiny of generic manufacturers following the event led authorities to uncover systematic manufacturing violations at NP. NP products that did not pass quality tests were

Ministry of Health, Labour and Welfare (MHLW) subsequently identified the resulting supply instability as a nationwide problem and introduced a series of policy measures to restore a stable supply of generic drugs (Ministry of Health, Labour and Welfare, 2023a). These developments underscore NP's role as a systemically important supplier in the Japanese generic pharmaceutical market. The misconduct was publicly reported on March 3, 2021, and a regulatory production-suspension order followed on March 5, 2021. We define March 3, 2021, as the onset of the NP shortage.

2.2 Drug Pricing Policies in Japan

Japan's universal health insurance covers all residents not receiving public assistance, who are enrolled in the National Health Insurance Program (NHIP). Financing comes primarily from social insurance premiums and taxes, with nearly 90% from public expenditure and slightly over 10% from individual contributions. Co-payment rates are set by age and income, leaving patients responsible for 10–30% of covered costs[3]. Coverage is comprehensive in scope. Nearly all pharmaceuticals for treatment approved in Japan are covered by the NHIP, and their prices are officially regulated.

These pharmaceuticals move from manufacturers to medical institutions via wholesalers. The officially set drug price ("yakka") is the retail price charged by these institutions, of which public health insurance typically covers 70–90%, leaving patients responsible only for the remaining coinsurance. It is revised biennially, with an additional interim annual adjustment starting in 2021. These revisions primarily align the official retail price with the competitive wholesale price at which hospitals and pharmacies purchase from wholesalers. In practice, wholesale prices are below the regulated price, and the gap constitutes a margin accruing to healthcare providers (Ministry of Health, Labour and

reprocessed through unapproved methods and distributed as if compliant; moreover, the manufacturing and quality-control systems were judged inadequate.

[3]In our data, the majority of patients are working-age and the copayment rate is 30%. More details are found in Section 4.

Welfare, 2025a).

Japanese generics are priced into three price tiers as follows (Ministry of Health, Labour and Welfare, 2023c). For newly listed generics, the official price is set at 50% of the corresponding brand-name price; if more than ten generics are listed simultaneously for an active ingredient, it is 40%. Where generics already exist, the new listing is priced at the lowest existing generic. Thereafter, listed generics undergo periodic retail price revisions, biennial with annual interim adjustments, which modify official prices. Because each product transacts at its own competitive wholesale price between providers and wholesalers, products sharing the same generic name can diverge in price. To standardize reimbursement, generics are grouped into three price tiers with weighted-average prices: a high tier at 50% or more of the maximum price, a middle tier at 30–50%, and a low tier at less than 30%. These rules are further adjusted for time since approval and other factors.

2.3 How Patients Obtain Prescription Drugs

In Japan, a patient first visits a clinic or hospital, where a physician issues a prescription, often written by generic (nonproprietary) name rather than for a specific product. The patient then presents the prescription at a pharmacy of their choice, which dispenses the medication[4]. The vast majority of pharmacies in Japan are privately owned, for-profit establishments[5].

At the pharmacy, the choice of which product to dispense rests largely with the pharmacist rather than the patient. Patients may express a preference between a brand-name

[4]Physicians may affect dispensing choices at the pharmacy, but this channel is limited in practice. Since the introduction of nonproprietary-name prescriptions in 2012, patients and pharmacists can generally choose between brand-name and generic products at pharmacies, and prescriptions that prohibit substitution account for less than 5% of prescriptions (Ministry of Health, Labour and Welfare, 2023d).

[5]According to the FY2020 Survey on Supply and Demand Trends of Pharmacists conducted by the Ministry of Health, Labour and Welfare, 56.6% of pharmacies were operated by joint-stock companies (kabushiki kaisha) and 29.7% by limited companies (yugen kaisha), together accounting for approximately 87% of pharmacies in the survey (Ministry of Health, Labour and Welfare, 2021, p. 89).

drug and a generic, whose price is typically about half that of the brand, and may in principle request a specific product; in practice, however, they receive whatever the pharmacy has in stock and generally choose neither the manufacturer nor the price tier of the generic dispensed. To contain inventory costs, a pharmacy typically stocks only one brand-name product and a single generic version from one manufacturer for each active ingredient[6]. Driven by government pay-for-performance incentives discussed in Section 2.4, dispensing has increasingly shifted toward generics.

This structure shapes how a supply disruption reaches patients. When the generic a pharmacy normally dispenses becomes unavailable, the pharmacist has two options. First, substitute the brand-name product, which typically increases patient costs. Second, dispense an alternative generic; depending on the price tiers of the unavailable and replacement products, patient costs may decrease, increase, or remain nearly unchanged. Which option a pharmacy takes, and, among generics, which alternative it selects, is therefore central to who bears the cost of a shortage, a point we formalize in Section 6.

2.4 Pay-for-Performance Incentives for Generic Drug Dispensing

To promote generic use and contain medical spending, the government introduced pay-for-performance incentives for pharmacies in 2008. According to the Ministry of Health, Labour and Welfare, the generic share (by quantity) rose from 32.5% in 2005 to 85.5% in 2024 (Ministry of Health, Labour and Welfare, 2025b). Beginning in 2010, incentives were tied to each pharmacy's generic-dispensing rate, calculated over the prior three months. Pharmacies meeting specified thresholds received a fixed add-on payment per

[6]For example, a pharmacy dispensing olmesartan, an angiotensin II receptor blocker (ARB) used to treat hypertension, may stock the originator brand (Olmetec) together with only one generic version (e.g., Nichi-Iko's olmesartan), rather than carrying generic products from multiple manufacturers.

prescription[7].

Subsequent fee-schedule revisions raised both thresholds and payments as generic-share targets increased. By 2020, the period analyzed, the thresholds were a generic-dispensing rate of $\geq$ 75% and $<$ 80% (150 JPY), $\geq$ 80% and $<$ 85% (220 JPY), and $\geq$ 85% (280 JPY) (Ministry of Health, Labour and Welfare, 2023b). These add-on fees are applied to all prescriptions dispensed at a pharmacy, including brand products. Given a contemporaneous fixed dispensing fee of 400 JPY per prescription, the scheme represented a sizable revenue component and created strong incentives to sustain high generic shares for pharmacies. As a result, over 70% of pharmacies had a generic share of at least 75% by 2020.

3 Empirical Methods

We first estimate the effects on patients' out-of-pocket spending and the types of drugs dispensed. To that end, we employ an event-study analysis using individual-level drug claims data. The treatment group consists of patients who used generic drugs manufactured by NP before the production suspension, whereas the control group consists of patients who used generic drugs manufactured by other manufacturers, as described in detail below. Consider a patient $i \in I$ receiving a drug in drug group $j \in J$ at pharmacy $k \in K$ in month $t \in T$. Each drug group consists of a brand-name product and generic products from one or more manufacturers. Once a patient receives a prescription, the pharmacy dispenses either a brand-name drug or a generic with the same active ingredient, dosage form, and strength.

To examine how the production suspension of Nichi-Iko Pharmaceutical (NP) affected

[7] In 2010, the categories and add-on payments were set as follows: 60 JPY per prescription for a share of $\geq$ 20% and $<$ 25%, 130 JPY for a share of $\geq$ 25% and $<$ 30%, and 170 JPY for a share of $\geq$ 30% (Ministry of Health, Labour and Welfare, 2010). A penalty for a low generic share was also introduced in 2018, initially targeting pharmacies with a generic share below 20%, and was extended in 2020 to pharmacies with a generic share below 40%. Such pharmacies were rare in our dataset.

drug choices and patient costs, our baseline event-study specification is

$$Y_{ijkt} = \sum_{\ell \neq -1} \delta_\ell D_{ij} \times \mathbf{1}\{E_t = \ell\} + \alpha_{ij} + \lambda_{jt} + \rho_{y(t)k} + \varepsilon_{ijkt}, \tag{1}$$

where Y_{ijkt} is the outcome variable, D_{ij} is a treatment indicator equal to one if patient i predominantly used NP products in drug group j before the suspension, and E_t denotes event time relative to the NP suspension. The omitted period is one month before the suspension, $E_t = -1$. The coefficients δ_ℓ trace the effect of exposure to NP products relative to the omitted pre-suspension period. The specification includes patient-by-drug-group fixed effects α_{ij}, month-by-drug-group fixed effects λ_{jt}, and fiscal-year-by-pharmacy fixed effects $\rho_{y(t)k}$, where $y(t)$ denotes the fiscal year containing month t. These interacted fixed effects absorb the corresponding patient, drug-group, month, and pharmacy fixed effects. We use two-way cluster-robust standard errors at the patient and pharmacy level.

Identification therefore comes from differential changes over time between NP-exposed and non-NP patient–drug-group cells within the same drug group and month, after controlling for pharmacy-specific shocks in each fiscal year. The coefficients δ_ℓ are constrained to be common across drug groups, so the specification pools comparisons conducted within drug groups rather than estimating a separate event study for each drug group. In this study, we examine two therapeutic classes that have large markets and require regular medication: dyslipidemia and hypertension. We define groups using the Japanese Pharmaceutical Efficacy Classification Code, which classifies products by pharmacologic action and therapeutic effect and is analogous to the U.S. NDC9 code. The system is hierarchical: the first digit denotes broad therapeutic areas, and three-digit codes index specific subclasses. We use the three-digit codes 214 for antihypertensives and 218 for dyslipidemia drugs.

Within each therapeutic class, we group products using the first nine characters of the YJ code, which identify the active ingredient, dosage form, and strength, and define this

unit as a *drug group* j. If a patient switches medications due to a drug shortage, the first choice of substitution is typically within the same drug group[8].

The treatment indicator D_{ij} is constructed based on patients' pre-suspension dispensing patterns. We first identify patients who frequently used generic drugs in each therapeutic class before the NP suspension and then classify them according to the manufacturer of the generic drugs they predominantly received. The treatment group comprises patients whose pre-suspension dispensations from NP account for at least 80% of their generic drug use within the same drug group. The control group comprises patients whose generic drug use consists of at least 80% from non-NP manufacturers. Thus, we compare patients receiving drugs from the supply-suspended manufacturer with patients receiving generic drugs from non-suspended manufacturers within the same *drug group*. Identification rests on a parallel-trends assumption: absent the NP suspension, outcomes for both groups would have followed similar trajectories.

We analyze four outcomes that capture two dimensions of the effects of the supply disruption: patient costs and the types of drugs dispensed. To measure patient costs, we first use the log of out-of-pocket spending per prescription. Out-of-pocket spending equals 30% of total drug spending for the prescription, reflecting the cost-sharing rate for the population in our analysis. We next use the log of the regulated drug price per dose to distinguish changes in spending arising from the price of the dispensed product from those arising from the quantity dispensed. To characterize product substitution, we use two indicator variables: one for whether the dispensed product is a brand-name drug, and one for whether it is a high-priced generic. We define the highest-priced generic as a

[8]The YJ code is a 12-digit standardized alphanumeric identifier used in Japan to classify and track individual pharmaceutical products uniquely. It encodes information about a drug's formulation, dosage form, manufacturer, and other distinguishing attributes, thereby facilitating regulatory oversight, data management, and product differentiation in the pharmaceutical market. For example, the first nine characters of the YJ code for Losartan Potassium Tablets 25 mg, one of the antihypertensive drugs, are "2149039F1." For distinct products in the same *drug group*, such as "NP" and "Sando," the complete codes are "2149039F1201" and "2149039F1287," respectively. Our drug group is defined based on the first nine digits of the YJ code.

generic product in the highest regulated price within the same drug group for each period.

4 Data

We use dispensing-claims data from the Rezult database provided by Japan System Techniques Co., Ltd. (JAST)[9]. The dataset spans June 2020–September 2024, covers roughly 9 million beneficiaries, and contains over 500 million claims[10]. Our analysis focuses on drugs used to treat hypertension and dyslipidemia, as defined above.

Each claim records a drug dispensed to one patient at a pharmacy[11]. We use the drug-group–patient–month–pharmacy level variation and estimate the event-study specification in equation (1). We timestamp the NP shock by its exact suspension date, which is feasible because dispensing dates are observed. Since the analytic data are monthly, we reassign the few pre-shock days within the shock month to the preceding month[12].

We construct the estimation sample in several steps. First, within hypertension and dyslipidemia, we restrict attention to drug groups in which NP supplied at least one product during the sample period. Second, we focus on patients who were regular generic-drug users before the suspension and classify exposure at the patient–drug-group level: patients are assigned to the NP-exposed group if NP products accounted for at least 80% of their pre-suspension generic use within the same drug group, and to the comparison group if their generic use was concentrated among non-NP manufacturers. Finally, we exclude drug groups whose product choice sets or pricing environments changed for reasons unrelated to the NP suspension[13].

[9] The database covers beneficiaries of Health Insurance Societies and Mutual Aid Associations, which primarily insure employees of relatively large firms and their dependents. Consequently, the study population is younger, more working-age, and higher income than the national average.

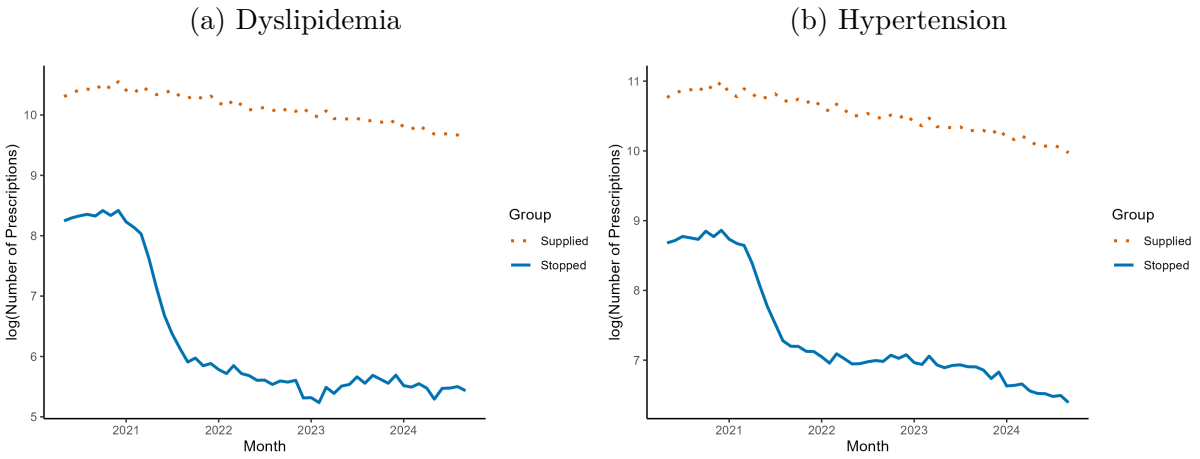
[10] We constructed the dataset after the state of emergency due to COVID-19 in April and May 2020.

[11] All personally identifiable information is anonymized; unique IDs are assigned.

[12] The misconduct was publicly reported on March 3, 2021. We assign March 1–3 to February and treat the remainder of March as the first post-suspension period.

[13] These exclusions remove drug groups with generic entry during the analysis period due to patent expiration, drug groups used only by one of the exposure groups or with too few exposed patients, drug groups

Figure 1: Prescription Counts for NP and Non-NP Products



Notes: The figure reports the log number of prescriptions for NP and non-NP products. The data are obtained from the Rezult database and cover June 2020 through September 2024. NP products are those manufactured by Nichi-Iko Pharmaceutical, and non-NP products are those manufactured by other firms.

4.1 Data Description

Figure 1 plots time paths of prescription counts in our sample by product category, contrasting NP products subject to supply suspension with non-NP products including branded drugs. Category panels are dyslipidemia (1a) and hypertension (1b). The patterns are consistent across categories: the number of prescriptions (in logs) for NP drugs ("Stopped") declines sharply beginning in March 2021, with no comparable drop for products from other manufacturers ("Supplied"). The timing coincides with public disclosure of the misconduct and the subsequent production-suspension order. Following NP's temporary production-suspension order, prescriptions for NP products remained at substantially reduced levels and did not recover their pre-suspension share.

The declining trend in prescription counts for both the "Stopped" and "Supplied" series reflects our sample restriction to patients who were prescribed these drugs before the suspension. As patients gradually exit the insurance claims database, the number of

with no brand-name dispensing over the relevant period, drug groups affected by changes in indications or dosage rules, and drug groups exhibiting abrupt price or cost discontinuities unrelated to the suspension. See details in Appendix Table A4

observed beneficiaries declines, mechanically reducing prescription counts.

Panel A in Table 1 shows statistics at the drug level. It shows a sharp fall in average brand prices—47% for dyslipidemia (from 63.1 to 33.3 JPY) and 24% for hypertension (from 79.1 to 59.8 JPY). These falls are mainly driven by regulated price revisions during the sample period. Generic prices declined modestly, 21% and 20%, respectively. In addition, before the shock, brand prices were, on average, 2.8 times higher in dyslipidemia and 2.9 times higher in hypertension. Supply expanded at the same time: the count of unique generics rose by about 6% for dyslipidemia and 4% for hypertension, with hypertension having a larger portfolio. While the number of unique drug groups remains constant, product density per group rises, with dyslipidemia showing a larger increase, from 16.7 to 17.8 items.

Panel B shows summary statistics of our study sample at the patient level. The average patients are male and in the late fifties, and the sample consists mostly of generic-drug users: the generic share exceeds 99% before the suspension in both therapeutic classes. After the shock, the generic share declines slightly, from 0.999 to 0.986 for dyslipidemia and from 0.999 to 0.987 for hypertension. In contrast, average out-of-pocket spending declines in both classes, from 314.7 to 250.6 JPY for dyslipidemia and from 339.4 to 290.4 JPY for hypertension, which may reflect changes in regulated drug prices. Our event-study estimates identify differential changes in drug choices and costs between patients exposed to NP products and comparable patients using non-NP generics. The number of observed patients declines in each class over the later period, consistent with the prescription-count patterns shown in Figure 1.

Appendix Table A3 shows the descriptive statistics at the pharmacy level. The estimation sample contains an average of 22,469 pharmacies per month, which is about 36% of the 62,828 operating nationwide in 2023 (Ministry of Health, Labour and Welfare, 2024). Across months, an average of 17.6% of pharmacies are in the no-add-on category, whereas 25.2% are in the highest add-on tier. The relatively large month-to-month variation in

Table 1: Descriptive statistics

	Dyslipidemia		Hypertension	
	Before	After	Before	After
Panel A: Drug characteristics				
Brand price	63.1	33.3	79.1	59.8
	(30.0)	(20.2)	(42.1)	(30.8)
	[17.5, 165.8]	[11.2, 146.6]	[9.2, 164.6]	[5.7, 163.9]
Generic price	22.2	17.6	27.1	21.8
	(9.46)	(7.28)	(13.7)	(10.6)
	[7.6, 69.2]	[7.6, 69.2]	[5.7, 88.7]	[5.7, 87.1]
Unique generic products	235	250	791	823
Unique drug groups	15	15	69	69
Products per drug group	16.7	17.8	12.2	12.9
	(7.91)	(8.50)	(8.66)	(8.57)
	[6, 28]	[7, 30]	[1, 30]	[2, 30]
Panel B: Patient characteristics				
Age	56.183	57.328	56.312	57.342
	(8.394)	(7.900)	(8.338)	(7.821)
Female ratio	0.407	0.395	0.318	0.308
	(0.491)	(0.489)	(0.466)	(0.462)
Generic share	0.999	0.986	0.999	0.987
	(0.028)	(0.119)	(0.036)	(0.114)
Out-of-pocket	314.650	250.618	339.409	290.444
	(235.840)	(201.528)	(251.677)	(230.391)
Unique patients	66,401	57,072	96,777	83,409

Notes: The data are obtained from prescription claims for two therapeutic classes after the sample selection. We exclude drug groups in which NP did not supply any products during the sample period, which runs from June 2020 to September 2024. Before indicates the period before the NP suspension, and After indicates the period on or after the suspension. Panel A reports drug-price and product-count statistics. Price statistics are calculated across prescriptions. Drug prices are measured in JPY per dose. Unique generic products and drug groups are distinct counts, while products per drug group summarizes the number of distinct products across drug groups. Panel B reports prescription-level patient characteristics and the number of distinct patients. Female ratio and generic share are proportions, and out-of-pocket payments are measured in JPY per prescription. Means are reported with standard deviations in parentheses, and square brackets report the minimum and maximum.

the highest tier indicates that the share of pharmacies qualifying for this tier changes substantially over time. Chain pharmacies account for an average of 32.5% of pharmacies in the sample.

5 Results

5.1 The Effect on Patients' Costs

Figure 2 reports the event-study estimates of the effect of the NP shock on patients' out-of-pocket spending and dispensed drug price.

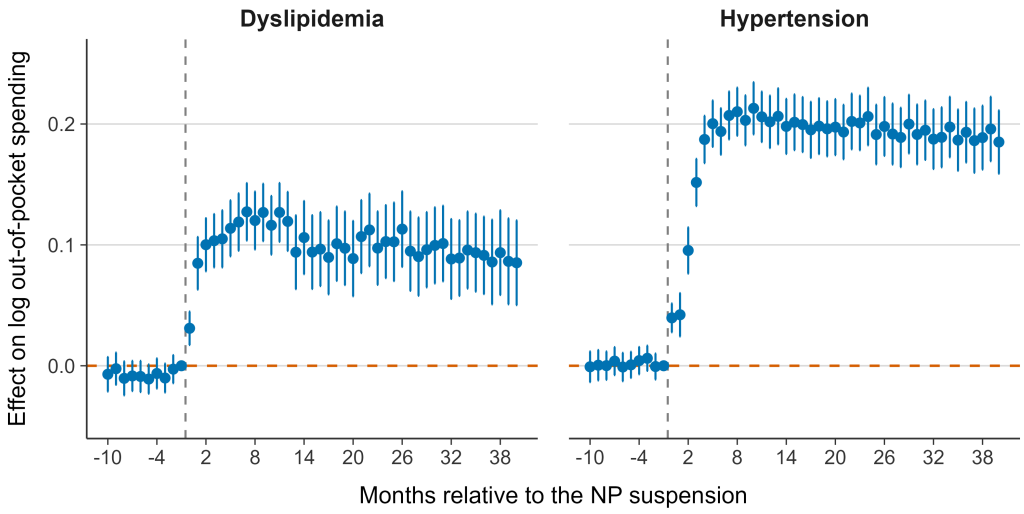
In panel (a), the estimates show an increase in out-of-pocket spending in both therapeutic classes, which remains elevated throughout the analysis period. For dyslipidemia, the effect is approximately 3.2% immediately after the suspension and 11.9% after one year, followed by a decline to 8.4% by the end of the observation period. For hypertension, it increases from 3.9% immediately after the suspension to 20.2% after one year and remains stable thereafter. These results indicate that the supply shock increased patients' out-of-pocket spending, and that this increase persisted for more than three years[14]

To understand how the increase in out-of-pocket spending occurs, panel (b) reports the estimates for the per-dose price of the dispensed drug. The pattern is qualitatively consistent with the patient's costs, showing that the NP shock increased the average price of the drugs dispensed in both therapeutic classes. For dyslipidemia, prices increase by approximately 1.4% immediately after the suspension and 5.6% after one year, followed by a decline to 4.1% by the end of the observation period. For hypertension, the corresponding estimates are 3.8% and 21.6%, and remain stable thereafter. These results indicate that the increase in patients' out-of-pocket spending is driven by substitution toward higher-priced drugs.

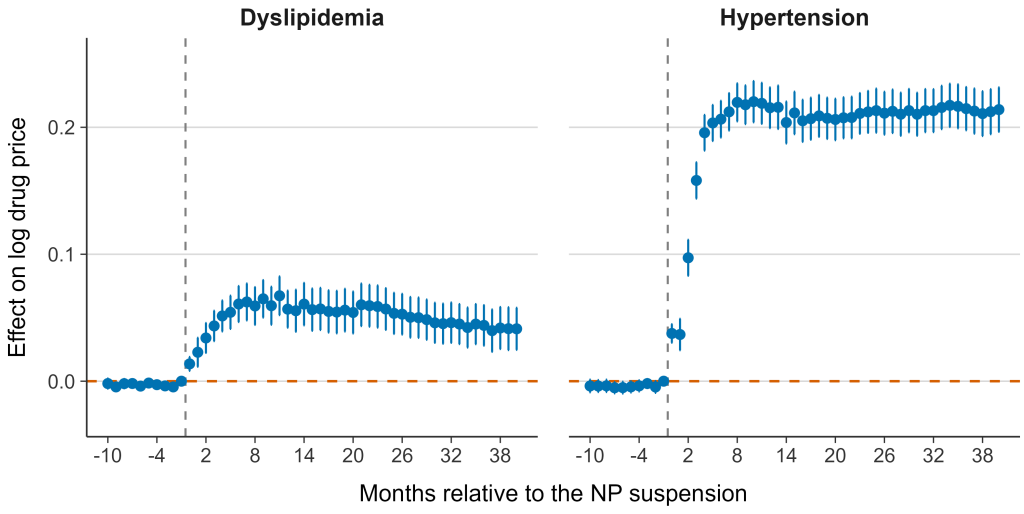
[14]The delayed onset of the increase likely reflects existing stocks held by pharmacies and wholesalers.

Figure 2: The effect on patients' costs

Panel (a): patient's copayment



Panel (b): dispensed drug price



Notes: The figure reports event-study estimates and 95% confidence intervals. In Panel (a), the outcome is the log of patient out-of-pocket spending per prescription. In Panel (b), the outcome is the log of the regulated drug price per dose. All specifications include patient-by-drug-group fixed effects, month-by-drug-group fixed effects, and fiscal-year-by-pharmacy fixed effects. Standard errors are two-way clustered at the patient and pharmacy levels. The vertical dashed line indicates the NP production suspension, and the omitted event month is one month before the suspension.

Appendix Figures A6 and A7 show that these results are robust to alternative definitions of treatment exposure, in which we change the pre-suspension generic-use threshold from 0.8 to either 0.7 or 0.9. Finally, one may be concerned that patients responded to the shock by switching pharmacies in search of available drugs, so that the estimates would partly reflect changes in where patients fill prescriptions rather than in what pharmacies dispense. Appendix Figure A11 shows little evidence of such switching in the short term: exposed patients are, if anything, less likely to visit a pharmacy they had not used before the suspension.

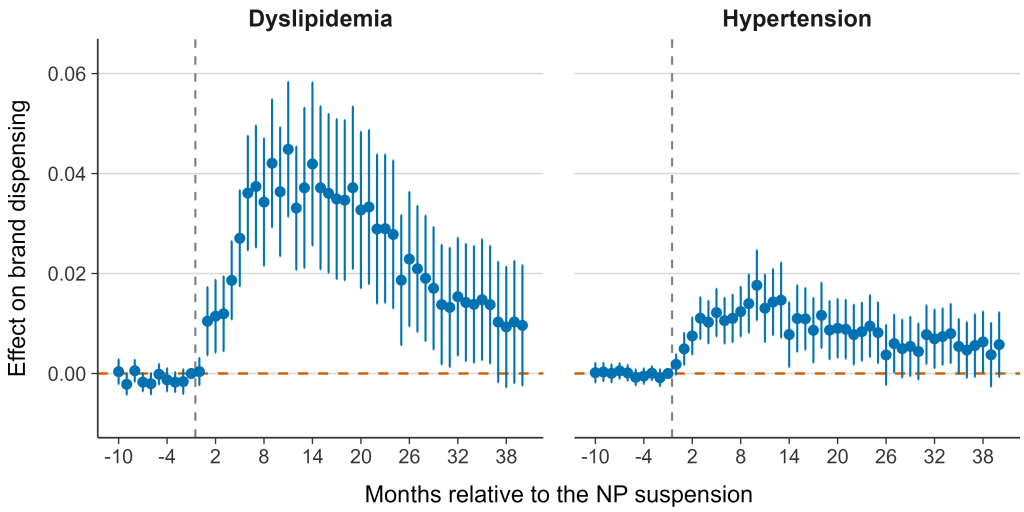
5.2 Drug Substitution

In Section 5.1, we found that patients' costs increased, and this increase was driven by a switch to higher-priced drugs. Mechanically, the replacement by a higher-priced drug could occur through a switch toward either a higher-priced generic or a brand product. Next, we decompose the increase in patients' costs. Note that the increased costs themselves do not represent a total welfare loss for patients, as some patients are willing to pay extra to receive an alternative generic (or brand) drug. To understand how we should interpret the increase in patients' costs, we provide a simple illustration of the drug market structure and the welfare loss from the increased costs in Appendix A.1.

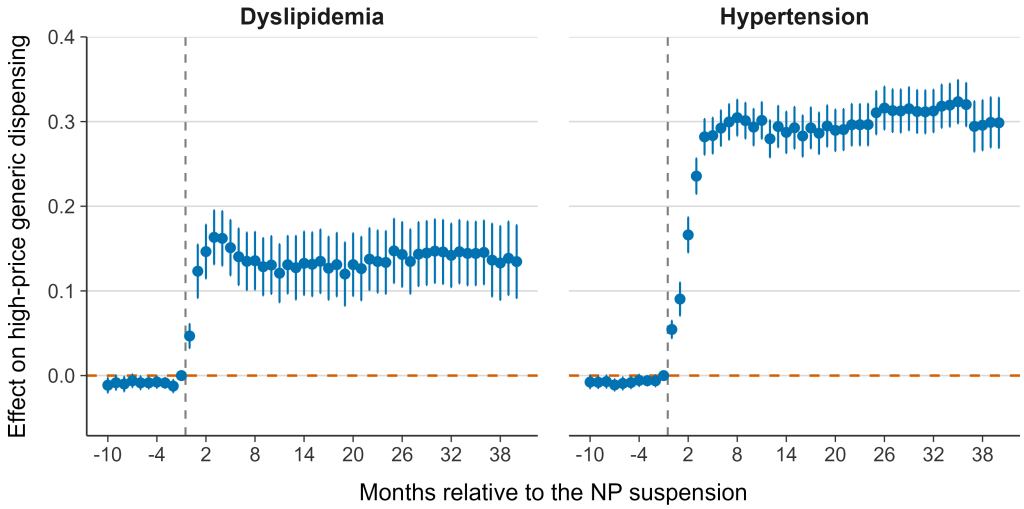
Figure 3 reports the event-study estimates of the effect of the supply shock on the probability of brand dispensing (Panel (a)) and the probability of highest-priced generic dispensing in a particular drug group (Panel (b)), using Equation (1). The two panels reveal two important points. First, the increase in patients' costs occurs through two different channels: one through a switch toward a brand drug, and the other through a switch toward a more expensive generic drug. The magnitude of the increase in highest-priced generic dispensing is far greater than the corresponding increase in brand dispensing, indicating that the switch itself was substantial along this second margin. Appendix Figure A3

Figure 3: The effect on drug substitution

Panel (a): brand share



Panel (b): highest-priced generic share

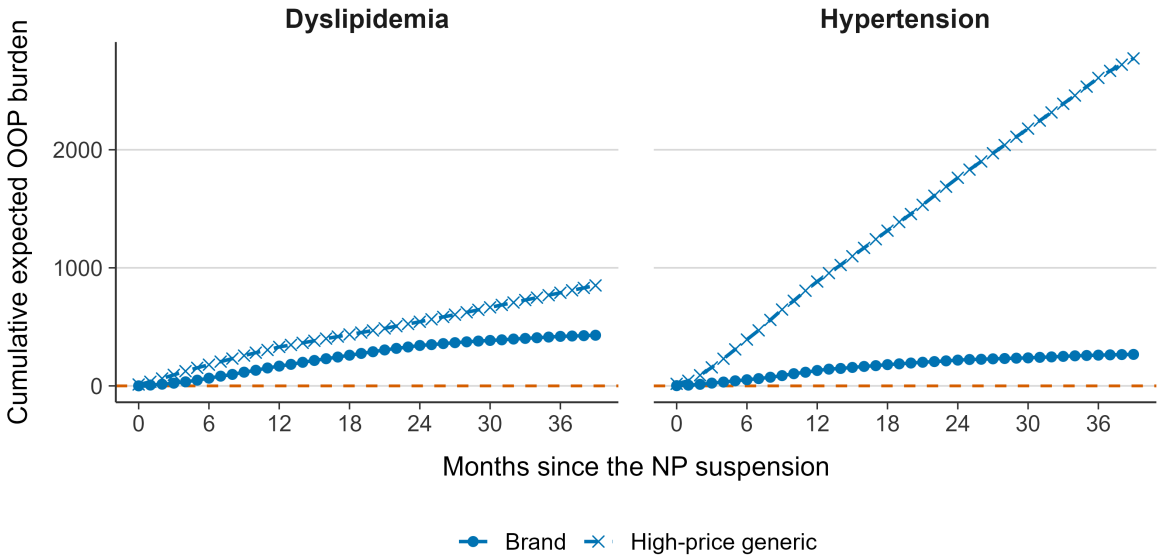


Notes: The figure reports event-study estimates and 95% confidence intervals. In Panel (a), the outcome is an indicator equal to one if the dispensed product is a brand-name drug. In Panel (b), the outcome is an indicator equal to one if the dispensed product is the highest-priced generic within its drug group and period. All specifications include patient-by-drug-group fixed effects, month-by-drug-group fixed effects, and fiscal-year-by-pharmacy fixed effects. Standard errors are two-way clustered at the patient and pharmacy levels. The vertical dashed line indicates the NP production suspension, and the omitted event month is one month before the suspension.

presents the pre-suspension distributions of regulated prices for NP and non-NP generic products. For dyslipidemia, the pre-suspension price distribution provides little evidence that the disappearance of lower-priced NP products mechanically drives the increase in highest-priced generic dispensing. For hypertension, however, NP generics were lower-priced and less likely to be in the highest-priced category before the suspension, implying that part of the post-suspension increase is mechanical. As shown in Appendix Table A1, the remaining response is nevertheless comparable in magnitude to the dyslipidemia response, supporting the interpretation that the suspension also induced substitution toward higher-priced generics. Second, the two panels display different post-suspension trajectories: the increase in brand-name dispensing is temporary, whereas the increase in highest-priced generic dispensing remains persistent. For brand-name dispensing, the estimates show a temporary rise that peaks 10 to 11 months after the supply shock. For dyslipidemia, the effect is about 0.1 percentage points immediately after the suspension, rises to 3.3 percentage points after one year, and then falls to 1.0 percentage point by the end of the observation period. For hypertension, the corresponding estimates are 0.2 and 1.4 percentage points, declining to 0.6 percentage points by the end of the period.For the probability of switching to the highest-price-category generic, by contrast, the effect appears immediately after the shock and persists, showing no reversion to the baseline. For dyslipidemia, it is about 4.7 percentage points immediately after the suspension and 13.1 percentage points after one year, remaining stable thereafter. For hypertension, the corresponding estimates are 5.5 and 28.1 percentage points, likewise stable thereafter.

The greater probability of switching to the highest-priced generic does not necessarily imply a larger cost burden, because the price gap between generics and brand-name drugs is larger than the price gap within generics. To quantify the relative importance of the two channels, Figure 4 reports a back-of-the-envelope calculation that combines the event-study estimates with observed out-of-pocket price gaps. For each event month and drug group, we compute the out-of-pocket gap relative to the lowest-priced generic, separately

Figure 4: Effect on Cumulative Patient Costs



Notes: This figure reports the cumulative contribution of substitution to branded drugs and to the highest-priced generic drugs to the increase in patients' out-of-pocket spending. For each event month and drug group, we calculate the average out-of-pocket price gap relative to the lowest-priced generic using prescription quantities as weights. The cumulative burden is obtained by multiplying these price gaps by the cumulative event-study estimates of branded and highest-priced generic dispensing shown in Figure 3.

for branded drugs and high-priced generics, average these gaps using prescription quantity as weights, and multiply them by the estimated increases in brand dispensing and high-priced generic dispensing in Figure 3. The results show that, despite the smaller within-generic price gap, the larger increase in high-priced generic dispensing makes within-generic substitution account for a larger share of the implied increase in patients' costs than the switch to branded drugs. Appendix A10 reports the corresponding month-specific contributions rather than their cumulative sums.

6 Pharmacy's Decision

The empirical results raise two questions. First, why do some patients receive brand-name products while others remain on generics? Second, why does brand-name substitution

decline over time, whereas substitution toward high-priced generics remains persistent? To answer these questions, we consider a simple model of pharmacy dispensing incentives. Pharmacies earn per-prescription profits and, if eligible, add-on payments tied to generic dispensing. When the original generic becomes unavailable, a pharmacy chooses in each period whether to dispense the brand-name product or procure an alternative generic at a procurement cost. The add-on payment shifts this choice toward generics but does not affect the choice among generics once a generic is already dispensed. This distinction allows the model to explain the cross-sectional choice between brand-name products and generics, while the structure of procurement costs explains the different time patterns between brand-name substitution and high-priced generic substitution[15].

6.1 Model

Consider pharmacies indexed by $k = 1, \ldots, n$ that maximize profits. For a given therapeutic class, we assume each pharmacy faces a stable demand $Q_k > 0$, given maintenance therapy for chronic disease. At any point in time, pharmacy k stocks one generic product and one brand-name product. Let $s_k \in \{0, 1\}$ indicate eligibility for the add-on payment tied to dispensing generics, with $s_k = 1$ if pharmacy k is eligible and $s_k = 0$ otherwise. The type s_k is treated as predetermined at the shock period. Before the supply shock, both the brand-name (b) and generic (g) are available. Because patients face lower costs for generics, they prefer g. Thus, in the pre-shock period, patients demand g, and pharmacies of both types dispense g. As Figure A11 shows, we assume that patient–pharmacy matches are fixed.

Once the supply is disturbed, it disrupts the original generic g. In each period $t \in \{1, \ldots, T\}$ after the shock, pharmacy k has two options. First, replace g with an

[15]Technically, a switch toward brand-name products could occur mechanically when no alternative generics are available. However, this channel is uncommon in both therapeutic classes.

alternative generic g', incurring a procurement cost $C_{k,t} \geq 0$ in that period[16]. Alternatively, switch to the brand-name b at zero additional cost, since the brand alternative is already stocked. Pharmacies' profits have two components: per-prescription margins and, when applicable, add-on payments. Let $\pi_{k,g'}$ and $\pi_{k,b}$ denote k's margins for dispensing a generic and a brand, respectively[17]. If pharmacy k qualifies for the add-on payment, it receives an additional amount Δ_S per prescription. If pharmacy k dispenses the brand-name product, it maintains eligibility for the add-on payment with probability $p_k \in [0, 1)$. For $s_k \in \{0, 1\}$, the pharmacy's decision in each period $t \in \{1, \ldots, T\}$ after the supply shock is therefore

$$\max\Big\{Q_k\,(\,\pi_{k,g'} + s_k\Delta_S) - C_{k,t},\; Q_k\,(\,\pi_{k,b} + s_k p_k\,\Delta_S\,)\Big\}.$$

Suppose pharmacy k is eligible for the add-on payment; then $s_k = 1$. If it procures the alternative generic g', it earns the generic margin and preserves the add-on payment. Therefore, its payoff is $Q_k(\pi_{k,g'} + \Delta_S) - C_{k,t}$. If it switches to the brand-name product b, it earns the brand margin but may lose eligibility for the add-on payment with continuation probability p_k. Its expected payoff is $Q_k(\pi_{k,b} + p_k\Delta_S)$. Thus, eligible pharmacies compare the brand margin advantage with the procurement cost of the alternative generic and the expected loss of add-on payments.

In contrast, when $s_k = 0$, pharmacy k is not eligible for the add-on payment. If it procures g', its payoff is $Q_k\pi_{k,g'} - C_{k,t}$. If it switches to b, its payoff is $Q_k\pi_{k,b}$. Thus, ineligible pharmacies choose between the alternative generic and the brand-name product solely by comparing product margins and the procurement cost of the alternative generic.

[16]The procurement cost accrues in each period in which the pharmacy dispenses the alternative generic: securing supply during a shortage requires ongoing search, negotiation, and order adjustment with wholesalers, rather than a one-time switching cost. We assume that each pharmacy cannot predict $C_{k,t}$ and solves the problem period by period, without dynamically optimizing its inventory.

[17]The dispensing profit for generic g at pharmacy k is $\pi_{k,g} = P^r_g - P^w_{k,g} \geq 0$, where P^r_g is the regulated retail price set by the official fee schedule and $P^w_{k,g}$ is the wholesale price at which pharmacy k acquires the drug. Similarly, $\pi_{k,b} = P^r_b - P^w_{k,b} \geq 0$ denotes the dispensing profit from the brand-name product.

6.1.1 Brand versus Generic Choice

Given the payoff maximization problem above, pharmacy k chooses the alternative generic g' rather than the brand-name product b if and only if

$$(\pi_{k,g'} - \pi_{k,b}) + s_k(1 - p_k)\Delta_S \geq c_{k,t}, \tag{2}$$

where $c_{k,t} = C_{k,t}/Q_k$ is the per-prescription procurement cost.

Equation (2) characterizes the brand-versus-generic choice in each period after the shortage. The generic-dispensing add-on raises the cutoff for choosing the generic because $s_k(1 - p_k)\Delta_S$ is positive only for add-on pharmacies. Conditional on the same product payoffs and procurement costs, add-on pharmacies are therefore more likely to procure an alternative generic rather than switch to the brand-name product.

We confirm this implication by estimating whether add-on pharmacies are more likely to continue dispensing generics after the supply shock:

$$Y_{ijkt} = \beta D_{ij} \times Post_t + \gamma\, Post_t \times AddOn_k + \delta D_{ij} \times Post_t \times AddOn_k + \mu_{ij} + \lambda_t + \varepsilon_{ijkt}, \tag{3}$$

where $AddOn_k$ equals one if pharmacy k received the generic-dispensing add-on before the suspension. When Y_{ijkt} is an indicator equal to one if the dispensed drug is generic, equation (2) predicts $\delta > 0$.

Table 2 reports the estimates. The results are consistent with the model prediction for hypertension: add-on pharmacies are more likely to continue dispensing generics after the supply shock. Although it does not reflect the causal relationship, this evidence supports the model's implication for the brand-versus-generic choice. Therefore, the add-on status is related to whether brand or generic is dispensed at each pharmacy.

Table 2: Difference in the impact on generic dispensing by pharmacy type

	Dyslipidemia		**Hypertension**	
	(1)	**(2)**	**(3)**	**(4)**
δ Add-on	0.003	-0.001	0.008	0.031
	(0.002)	(0.002)	(0.002)	(0.015)
Individual FE	×	×	×	×
Month FE	×	×	×	×
Pharmacy FE	×	×	×	×
Drug-group FE	×	×	×	×
FY × Pharmacy FE	×	×	×	×
Month × Drug-group FE		×		×
Observations	1,379,993	1,379,993	2,092,062	2,092,062
N patients	64,590	64,590	93,400	93,400

Notes: The table reports the coefficient on the interaction between the treatment indicator, the post-suspension indicator, and a dummy variable indicating whether the pharmacy received the generic-dispensing add-on before the suspension. The outcome variable is a binary variable equal to one if the dispensed drug is generic and zero otherwise. Odd-numbered columns control for individual and month fixed effects. Even-numbered columns additionally include pharmacy fixed effects and their interaction with fiscal-year dummies.

6.2 Within-Generic Incentives

The preceding subsection explains why add-on pharmacies are more likely to remain with generics rather than to return to the original brand drugs. Thus, we now turn to the second empirical pattern in Figure 3; brand-name substitution is temporary, whereas substitution toward high-priced generics is persistent.

6.2.1 The Role of the Generic-Dispensing Add-on

Suppose that, after adopting a high-priced generic g^H, pharmacy k can switch to a lower-priced generic $g^L \in \mathcal{G}^L_{k,t}$, where $\mathcal{G}^L_{k,t}$ denotes the set of lower-priced generics available to pharmacy k in period t. For simplicity, we assume two price categories. Let $c^H_{k,t}$ and $c^L_{k,t}$ denote the per-prescription procurement costs of dispensing high-priced and lower-priced generics, respectively, measured relative to dispensing the brand-name product. Define

the procurement-cost differential as $\Delta c_{k,t} = c^L_{k,t} - c^H_{k,t}$. This differential may take either sign[18].

Continuing to dispense g^H yields the per-prescription payoff $\pi_{k,g^H} + s_k\Delta_S - c^H_{k,t}$, while switching to a lower-priced generic g^L yields $\pi_{k,g^L} + s_k\Delta_S - c^L_{k,t}$. Pharmacy k receives $s_k\Delta_S$ in both cases, so the add-on cancels out. The level of generic procurement costs also does not matter for the within-generic comparison; only the cost differential $c^L_{k,t} - c^H_{k,t}$ matters. Switching from g^H to the best available lower-priced generic, therefore, requires

$$\max_{g^L \in \mathcal{G}^L_{k,t}} \left\{ \pi_{k,g^L} - \pi_{k,g^H} \right\} \geq \Delta c_{k,t}.$$

Therefore, the add-on does not affect the choice among generics. A similar logic applies to procurement costs: their level affects the brand-versus-generic condition in equation (2), but only the differential $\Delta c_{k,t}$ affects the choice among generics. Suppose that the procurement-cost levels $c^H_{k,t}$ and $c^L_{k,t}$ decline over time as the supply of alternative generics recovers. As the procurement cost of the adopted generic falls below the cutoff in equation (2), pharmacies that switched to the brand-name product return to a generic. This recovery, however, does not induce a return from g^H to a lower-priced generic, which occurs only if the margin gain from lower-priced generics exceeds the cost differential $\Delta c_{k,t}$. The persistence of high-priced generic substitution in Figure 3 is consistent with the pattern if this condition fails for most pharmacies[19]. Thus, these financial incentives explains why brand-name substitution is temporary whereas substitution toward high-priced generics is persistent.

This is the limitation of the existing add-on. It encourages pharmacies to stay within

[18] For expositional simplicity, the model groups generics into high-priced and lower-priced products. Extending the choice set to three price categories does not change the result. The differential $\Delta c_{k,t}$ excludes differences in wholesale acquisition prices across categories because those differences are already incorporated into the dispensing profits $\pi_{k,m}$.

[19] The condition can fail because lower-priced generics generate smaller per-prescription payoffs ($\pi_{k,g^L} - \pi_{k,g^H} \leq 0$), or because the cost differential $\Delta c_{k,t}$ is positive, or both. Either channel implies that the recovery of the common procurement cost leaves the within-generic choice unchanged.

generics, but it does not encourage them to choose lower-priced generics once a generic is already dispensed. A policy that rewards generic dispensing can therefore maintain generic use while still failing to prevent shortage-induced increases in patient costs.

6.2.2 Within-Generic Pay-for-Performance Incentives

In this subsection, we consider a hypothetical policy to promote substitution to lower-priced generics. Specifically, we consider a pay-for-performance incentive τ paid per prescription when pharmacy k dispenses a lower-priced generic $g^L \in \mathcal{G}^L_{k,t}$. With this incentive, switching from g^H occurs if

$$\max_{g^L \in \mathcal{G}^L_{k,t}} \left\{ \pi_{k,g^L} - \pi_{k,g^H} \right\} + \tau \geq \Delta c_{k,t}.$$

The minimum incentive required to induce switching is therefore

$$\tau^*_{k,t} = \max \left\{ \Delta c_{k,t} - \max_{g^L \in \mathcal{G}^L_{k,t}} \left\{ \pi_{k,g^L} - \pi_{k,g^H} \right\}, 0 \right\}.$$

The required incentive compensates for the procurement-cost differential of lower-priced generics, net of the margin difference between the two price categories. In contrast to the existing add-on, which operates only on the brand-versus-generic margin in equation (2), the incentive τ operates directly on the within-generic margin.

We implement this calculation using product- and period-specific implied dispensing profits estimated from the April 2022, April 2023, and April 2024 price revisions. For each policy-relevant high-priced generic prescription, we calculate the incentive required to match its implied profit using the most profitable product among the lowest-priced alternatives in the same drug group and month. We report both the average across observed post-period prescriptions and a targeted estimate for prescriptions induced by the shortage, obtained as the ratio of the corresponding DID estimates. Appendix C

provides the details.

Table 3: Within-Generic Pay-for-Performance Incentives

	Dyslipidemia		Hypertension		All	
	Average	Targeted	Average	Targeted	Average	Targeted
Incentive τ	24	28	101	75	77	60
Patient						
Gross drug-cost saving	60	130	180	256	142	215
Patient contribution to incentive	7	8	30	23	23	18
Net patient saving	53	121	150	234	119	197
Insurer						
Gross drug-cost saving	144	306	431	608	338	508
Insurer contribution to incentive	17	20	71	53	54	42
Net fiscal saving	127	286	360	555	285	466
Medical Spending						
Gross drug-spending reduction	205	436	611	864	480	722
Net medical-spending reduction	180	407	510	789	404	663
Incentive-to-gross-saving ratio	11.8%	6.4%	16.6%	8.7%	15.9%	8.2%
Break-even Δc	180	407	510	789	404	663

Notes: This table shows the estimated incentives τ and its cost-effectiveness for both therapeutic classes. Monetary values are JPY per prescription. Incentive τ is treated as an insured dispensing fee. Patients pay 30% of the incentive and insurers pay the remainder. Net patient saving and net fiscal saving deduct the respective incentive contributions from gross drug-cost savings. Net medical-spending reduction equals the reduction in drug spending less the incentive payment. Break-even Δc is the additional procurement-cost differential that reduces net medical-spending savings to zero.

Table 3 reports the results. The Average column evaluates the incentive over all observed policy-relevant high-priced generic prescriptions in the post period, while the Targeted column evaluates it over the prescriptions induced by the shortage, using the ratio of DID estimates. Comparing the two columns isolates how shortage-induced substitution differs from the average policy-relevant prescription along two dimensions: the required incentive (τ) and the gross drug-spending reduction (S^T).

For dyslipidemia, the required incentive is modestly higher for shortage-induced prescriptions than for the average prescription (28 JPY versus 24 JPY), while the gross drug-spending reduction more than doubles (436 JPY versus 205 JPY). For hypertension, the required incentive is instead lower for shortage-induced prescriptions (75 JPY versus 101

JPY), while the gross drug-spending reduction again rises substantially (864 JPY versus 611 JPY). In both therapeutic classes, therefore, the gross saving from shortage-induced substitution grows by more than the required incentive, so the incentive-to-gross-saving ratio falls in both cases—from 11.8 to 6.4 percent for dyslipidemia and from 16.6 to 8.7 percent for hypertension—even though τ itself moves in opposite directions across the two classes. This opposite movement reflects the composition of shortage-induced substitution across products with different pharmacy profit differentials. The decomposition in Appendix Table A2 shows that, relative to the average policy-relevant prescription, shortage-induced substitution in dyslipidemia is concentrated in products with a *larger* raw pharmacy profit differential ($\pi_{k,g^H} - \pi_{k,g^L}$), whereas in hypertension it is concentrated in products with a *smaller* one. Because the required incentive is the positive part of this differential rather than the differential itself, it need not track the raw differential one-for-one[20]. In both classes, however, the gross drug-spending reduction rises by more than the required incentive, so the ratio falls regardless of the direction in which τ moves.

Pooling the two therapeutic classes, the average required incentive is 77 JPY, falling to 60 JPY in the targeted calculation, while the gross drug-spending reduction rises from 480 JPY to 722 JPY. The incentive-to-gross-saving ratio therefore falls from 15.9 to 8.2 percent. After accounting for the patient and insurer contributions to the incentive, net savings in the targeted calculation are 197 JPY for patients, 466 JPY for insurers, and 663 JPY in total medical spending—each roughly 60 percent larger than the corresponding Average figures. Finally, adding the break-even Δc to the required incentive at $\Delta c = 0$ gives the maximum total incentive that can be paid without increasing net medical spending; this upper bound equals the gross drug-spending reduction and is substantially larger, in both the Average and Targeted calculations, than the incentive itself.

[20] In dyslipidemia, the shortage-induced prescriptions lie almost entirely on the incentive-requiring side, so the truncation that inflates the required incentive above the raw differential for the average prescription largely disappears; the required incentive therefore rises far less than the raw differential even as that differential widens.

7 The Effect on Patients' Discontinuation and Adherence

Supply disruptions may affect both whether patients continue treatment and how consistently they obtain medication. We therefore estimate patient-level regressions separately by outcome, therapeutic class, and post-suspension period. We consider two patient-level channels: patients may discontinue medication following the supply shock, or they may reduce adherence while continuing treatment. Specifically, for each patient i, we estimate

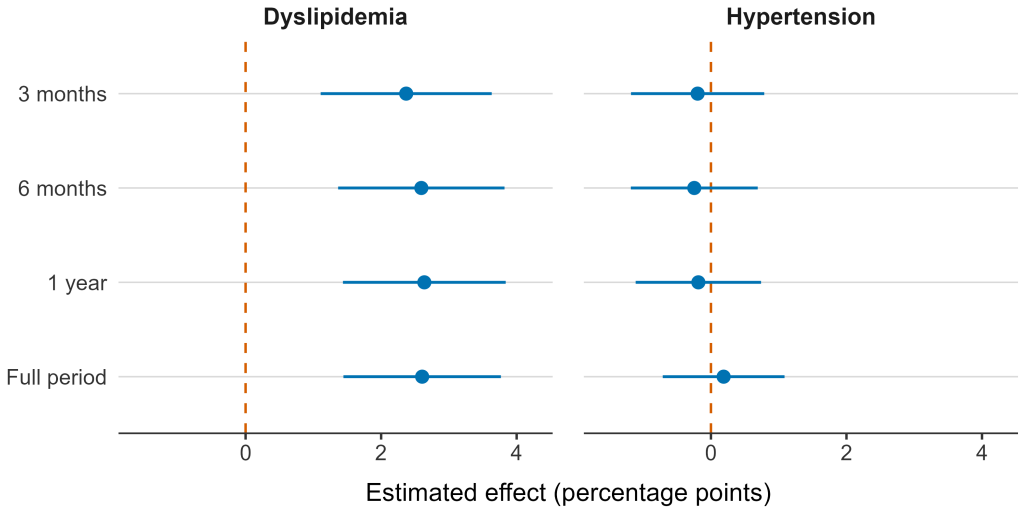
$$Y_i = \alpha + \beta D_i + X_i'\gamma + \varepsilon_i, \tag{4}$$

where D_i indicates whether patient i used NP products before the suspension within the corresponding therapeutic class, X_i includes the patient's pre-suspension age and sex, and ε_i is an error term. The dependent variable Y_i is one of two outcomes: discontinuation and adherence. Standard errors are clustered at the level of the pre-suspension pharmacy, defined as the pharmacy most frequently used by the patient before the suspension.

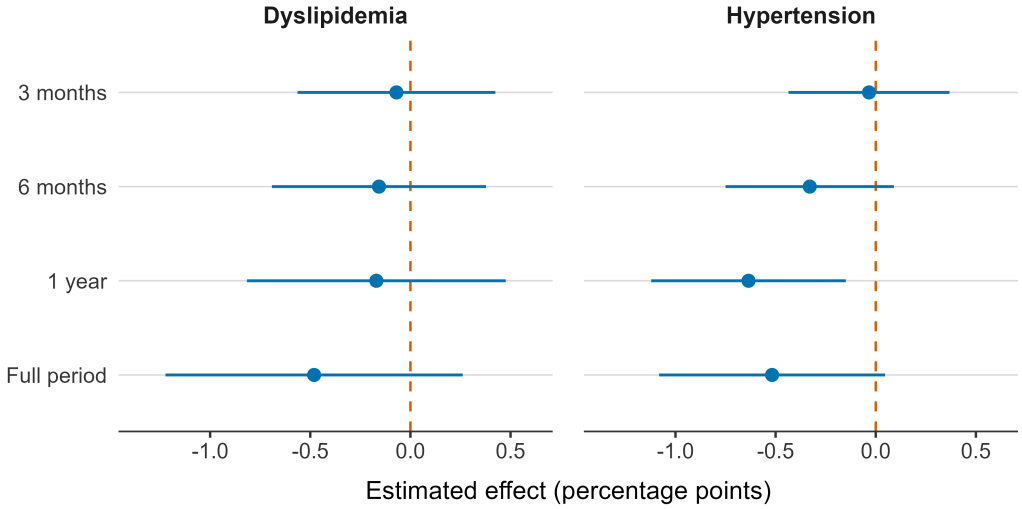
To measure discontinuation and adherence, we consider four post-suspension periods: three months, six months, one year, and the full observation period. Discontinuation measures whether patients use medication at all after the shock. A patient is classified as discontinuing medication if no dispensing claim is observed in the sample during the corresponding post-suspension period. Adherence measures how consistently patients use medication among those who continue to receive at least one dispensing after the shock. For each patient, we compute the share of months with any dispensing claim before the suspension and the corresponding share during the post-suspension period. The adherence outcome is defined as the change in this dispensing rate from before to after the suspension. Thus, discontinuation and adherence capture the extensive and intensive margins of medication access due to the suspension shock, respectively.

Figure 5: The Effect on Treatment Continuity

Panel (a): treatment discontinuation



Panel (b): medication adherence



Notes: Panel (a) reports estimates of the effect of the NP shock on treatment discontinuation, and Panel (b) reports estimates of the effect on changes in medication adherence. Each point represents the estimated coefficient for a given post-suspension period and therapeutic class. Horizontal bars report 95% confidence intervals based on standard errors clustered at the pre-suspension pharmacy level.

Figure 5 reports the estimated effects of the supply shock on discontinuation and adherence. The results in Panel (a) for discontinuation differ across the two therapeutic classes. The estimates show a clear increase in discontinuation for dyslipidemia. The effect appears within the first three months after the suspension and remains positive through the full observation period, with the long-run effect around two percentage points. In contrast, the estimates for hypertension are small and close to zero across periods.

Panel (b) reports the estimated effects on adherence. The estimates tend to be negative for both therapeutic classes over a longer period, indicating that patients exposed to the NP shock experienced larger declines in dispensing frequency than those who were not. For hypertension, the decline becomes visible by one year and persists through the full observation period. Overall, these findings suggest that the NP shock affected not only the extensive margin but also the intensive margin of medication access.

8 Conclusions

Supply disruptions are a pervasive feature of modern economies, and how their costs are allocated depends critically on whether prices can adjust. In flexible-price markets, scarcity is largely resolved through rising prices; in regulated markets, such as housing, agricultural goods, health care, prices cannot clear shortages, and the disruption must instead be absorbed through non-price mechanisms that determine who ultimately bears its cost. These non-price costs are especially hard to quantify when regulated pricing is compounded by delegated product choice: when a single regulated price covers a category of close but differently priced substitutes, and the party selecting among them is not the end consumer but an intermediary acting on the consumer's behalf—a pharmacist, physician, or contractor. Under these conditions, a supply disruption need not appear as a price increase; it can surface as a shift in which product is provided, with cost consequences that depend on the intermediary's choice rather than the consumer's. This

paper quantifies these consequences in the Japanese generic pharmaceutical market, where prices are set by regulation, pharmacists mediate between prescriptions and dispensed products, and a large, government-ordered production suspension provides an exogenous shock with which to identify the effects of shortages on patients.

Our results show that patients' out-of-pocket spending per prescription increased by up to 21% following the shortage, even though the shortage itself did not generate market-clearing price increases. This finding illustrates how the two features of this market translate a supply disruption into a cost increase: because clinically equivalent products are sold at multiple regulated prices, and because pharmacists rather than patients select among them, a shortage can raise costs simply by shifting which product is dispensed. We show that this shift operates through two distinct channels: patients were more likely to be switched from generics to brand drugs, and, perhaps more surprisingly, patients remaining on generics were more likely to be switched to a higher-priced generic version of the same drug. This latter channel, a shift within the generic segment itself, has received little attention in the prior literature, and we find that it accounts for a larger share of the overall cost increase than the switch to brand drugs because far more prescriptions move toward higher-priced generics and because that shift persists, consistent with the generic-dispensing add-on dropping out of the choice among generics.

What determines a pharmacy's choice between these two paths—switching patients to a higher-priced generic or to a brand drug? Our findings suggest that it depends on financial incentives pharmacies receive for dispensing generics. The likelihood of switching to brand-name drugs was smaller among patients obtaining medications from pharmacies with stronger generic-dispensing incentives than among those without, consistent with the program's incentive to remain within the generic segment. We rationalize this heterogeneous response to shortages using a simple economic model of pharmacy profits that emphasizes the procurement costs pharmacies incur when sourcing alternative products during shortages. Given the rise in patient costs, we further analyze whether shortages af-

fect medication discontinuation and adherence, given the rise in patient costs. Our results indicate that patients were more likely to discontinue treatment and less likely to adhere to their medication following shortages, with the magnitude and statistical significance of these effects varying across therapeutic classes and post-suspension horizons, suggesting that the increase in out-of-pocket spending captures only part of the true welfare cost of drug shortages. Such disruptions in treatment may have serious health consequences even when they are relatively short-lived. Chandra et al. (2024) show that temporary medication interruptions can substantially increase mortality, particularly when patients discontinue high-value drugs, and that patients may not fully recognize these risks. This lack of awareness makes disruptions induced by drug shortages particularly concerning, as patients may underestimate the health consequences of even temporary interruptions in treatment.

Taken together, these results carry two distinct policy implications. First, the finding that pharmacies without financial incentives are more likely to substitute brand-name drugs for generics underscores the importance of incentive schemes such as the generic-dispensing add-on payment: absent such incentives, shortages can undo part of the cost savings that generic-substitution policies are designed to achieve. Second, because the shift toward higher-priced generics contributes more to the rise in patient costs than the shift toward brand-name drugs, current schemes, effective at keeping patients within the generic market, are not designed to encourage substitution toward the least expensive generic alternative. Closing this gap may require a more finely targeted incentive that rewards pharmacies not merely for dispensing generics in general, but specifically for dispensing lower-priced ones. Our back-of-the-envelope calculations suggest such an incentive would more than pay for itself: an add-on of approximately JPY 60 per shortage-induced high-priced-generic prescription would reduce patient out-of-pocket spending by JPY 197 and total medical spending by JPY 663, net of the incentive cost.

More broadly, our findings suggest that in markets where fixed pricing creates room for

horizontal substitution and intermediaries mediate product choice, incentive design must account not only for whether the intermediary steers consumers toward a broad product category, such as "a generic," but also for which specific product within that category is ultimately chosen. Failing to address this margin may leave a substantial share of the welfare cost of supply disruptions unaddressed, even under otherwise well-designed incentive schemes.

References

Agarwal, N., I. Ashlagi, M. A. Rees, P. Somaini, and D. Waldinger (2021): "Equilibrium Allocations Under Alternative Waitlist Designs: Evidence From Deceased Donor Kidneys," *Econometrica*, 89, 37–76.

Barrot, J.-N. and J. Sauvagnat (2016): "Input Specificity and the Propagation of Idiosyncratic Shocks in Production Networks," *Quarterly Journal of Economics*, 131, 1543–1592.

Boehm, C. E., A. Flaaen, and N. Pandalai-Nayar (2019): "Input Linkages and the Transmission of Shocks: Firm-Level Evidence from the 2011 Tōhoku Earthquake," *Review of Economics and Statistics*, 101, 60–75.

Brekke, K. R., T. H. Holmås, and O. R. Straume (2013): "Margins and Market Shares: Pharmacy Incentives for Generic Substitution," *European Economic Review*, 61, 116–131.

Bulow, J. and P. Klemperer (2012): "Regulated Prices, Rent Seeking, and Consumer Surplus," *Journal of Political Economy*, 120, 160–186.

Carvalho, V. M., M. Nirei, Y. U. Saito, and A. Tahbaz-Salehi (2021): "Supply Chain Disruptions: Evidence from the Great East Japan Earthquake," *Quarterly Journal of Economics*, 136, 1255–1321.

Chalkley, M. and C. Tilley (2005): "The Existence and Nature of Physician Agency: Evidence of Stinting from the British National Health Service," *Journal of Economics & Management Strategy*, 14, 647–664.

Chan, D. C., M. Gentzkow, and C. Yu (2022): "Selection with Variation in Diagnostic Skill: Evidence from Radiologists," *Quarterly Journal of Economics*, 137, 729–783.

Chandra, A., E. Flack, and Z. Obermeyer (2024): "The Health Costs of Cost Sharing," *The Quarterly Journal of Economics*, 139, 2037–2082.

Clemens, J. and J. D. Gottlieb (2014): "Do physicians' financial incentives affect medical treatment and patient health?" *American Economic Review*, 104, 1320–1349.

Conti, R. M. and M. E. Wosińska (2025): "The Economics of Generic Drug Shortages: The Limits of Competition," *Journal of Economic Perspectives*, 39, 79–102.

Dubois, P., G. Majewska, and V. Reig (2023): "Drug shortages: empirical evidence from France," *working paper*.

European Commission (2025): "Commission Staff Working Document: Summarising Evidence Supporting the Legislative Proposal Laying a Framework for Strengthening the Availability and Security of Supply of Critical Medicinal Products as Well as the Availability of, and Accessibility of, Medicines of Common Interest and Amending Regulation (EU) 2024/795," Tech. Rep. SWD(2025) 263 final, European Commission, Brussels.

Frech, III, H. E. and W. C. Lee (1987): "The Welfare Cost of Rationing-by-Queuing Across Markets: Theory and Estimates from the U.S. Gasoline Crises," *Quarterly Journal of Economics*, 102, 97–108.

Glaeser, E. L. and E. F. P. Luttmer (2003): "The Misallocation of Housing Under Rent Control," *American Economic Review*, 93, 1027–1046.

Gravelle, H. and L. Siciliani (2008): "Optimal Quality, Waits and Charges in Health Insurance," *Journal of Health Economics*, 27, 663–674.

Huitfeldt, I., V. Marone, and D. Waldinger (forthcoming): "Designing Dynamic Reassignment Mechanisms: Evidence from GP Allocation," *American Economic Review*, forthcoming.

Iizuka, T. (2012): "Physician Agency and Adoption of Generic Pharmaceuticals," *American Economic Review*, 102, 2826–2858.

Ito, Y., K. Hara, and Y. Kobayashi (2020): "The Effect of Inertia on Brand-name versus Generic Drug Choices," *Journal of Economic Behavior & Organization*, 172, 364–379.

Janssen, A. and D. Granlund (2023): "The Importance of the First Generic Substitution: Evidence from Sweden," *Journal of Economic Behavior & Organization*, 213, 1–25.

Kakehi, H. and R. Nakajima (2026): "Role of Pharmacists in Generic Pharmaceutical Adoption," Working paper.

McGuire, T. G. (2000): "Physician Agency," in *Handbook of Health Economics*, ed. by A. J. Culyer and J. P. Newhouse, Elsevier, vol. 1 of *Handbook of Health Economics*, 461–536.

Ministry of Health, Labour and Welfare (2010): "Explanatory Materials for the Medical Fee Revision in the 22nd Year of Heisei," https://www.mhlw.go.jp/bunya/iryouhoken/iryouhoken12/setumei.html (in Japanese).

——— (2021): "Survey on Supply and Demand Trends of Pharmacists," Tech. rep., Ministry of Health, Labour and Welfare, fY2020 survey. In Japanese.

——— (2023a): "Current Status of the Generic Drug Industry," .

——— (2023b): "Explanatory Materials for the Medical Fee Revision in the Fourth Year of Reiwa," https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000196352_00008.html (in Japanese).

——— (2023c): "On the Drug Price Revision for Fiscal Year 2024 (Reiwa 6)," https://www.mhlw.go.jp/stf/shingi2/0000212451_00068.html (in Japanese).

——— (2023d): "Survey on the Use of Generic Drugs," .

——— (2024): "Report on Public Health Administration and Services: Summary Report, FY2023 [5]," Tech. rep., Ministry of Health, Labour and Welfare, Japan, https://www.mhlw.go.jp/toukei/saikin/hw/eisei/23/ (in Japanese).

——— (2025a): "Drug Price Survey," https://www.mhlw.go.jp/toukei/list/157-1b.html (in Japanese).

——— (2025b): "Survey on the Use of Generic Drugs," https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/iryou/kouhatu-iyaku/04.html (in Japanese).

Nichi-Iko Co., Ltd. (2021): "Administrative Disposition Imposed on Nichi-Iko Co., Ltd. by the Governor of Toyama Prefecture," Press release; Available at: https://www.nichiiko.co.jp/company/press/detail/5087/1301/4541_20210303_01.pdf (in Japanese).

Nichi-Iko Pharmaceutical Co., Ltd. (2021): "Consolidated Financial Results for the Fiscal Year Ended March 31, 2021," .

Nichi-Iko Pharmaceutical Co., Ltd. and Eisai Co., Ltd. (2019): "Nichi-Iko Pharmaceutical and Eisai Agree on Strategic Business Alliance," .

Ridley, D. B., X. Bei, and E. B. Liebman (2016): "No shot: US vaccine prices and shortages," *Health Affairs*, 35, 235–241.

Stomberg, C. (2018): "Drug Shortages, Pricing, and Regulatory Activity," in *Measuring and Modeling Health Care Costs*, ed. by A. Aizcorbe, C. Baker, E. R. Berndt, and D. M. Cutler, University of Chicago Press, 323–348.

Yurukoglu, A., E. Liebman, and D. B. Ridley (2017): "The role of government reimbursement in drug shortages," *Am. Econ. J. Econ. Policy*, 9, 348–382.

Appendix

A Supply Disruptions and Patient Cost Burden

A.1 Simple Graphical Framework

In this section, we present a simple graphical framework to illustrate how supply disruptions are transmitted to patients under different institutional environments. Supply disruptions can affect patients through multiple channels. In standard competitive markets, shortages are typically resolved through price adjustment. In regulated healthcare markets, however, prices cannot fully adjust, leaving non-price mechanisms to allocate scarce supply. In pharmaceutical markets, one important non-price mechanism is substitution from a drug affected by supply disruptions (hereafter, the affected drug) to a higher-price alternative drug.

A central feature of this framework is that patients do not directly choose between the affected drug and the higher-price drug. Instead, pharmacists act as supply-side intermediaries who make substitution decisions on patients' behalf. This institutional structure means that market outcomes are shaped not only by patient preferences but also by pharmacist incentives, a feature that is reflected throughout the graphical framework. Figure A1 illustrates the basic intuition underlying our framework. The left panel shows the affected-drug market. The government-ordered production suspensions shifted the supply of the affected drug leftward from S_0 to S_1. Because the affected drug's price is regulated at P_A, the market cannot clear through price increases. Instead, the available quantity falls from Q_0 to Q', generating a shortage of magnitude $Q_0 - Q'$. Patients who can no longer obtain the affected drug lose consumer surplus, represented by the shaded triangle.

The right panel shows the higher-price drug market. The upward-sloping curve MC is the marginal cost of producing and dispensing the higher-price drug and is unaffected by the shock. Before the disruption, we assume that this market is at the point E_1^*, where

the demand curve intersects the marginal cost curve exactly at the regulated price P_H: at the baseline quantity Q_0, marginal willingness to pay equals marginal cost. This is a deliberate abstraction rather than a description of equilibrium behavior: it takes the pre-shock allocation in the higher-price drug market to be free of pre-existing distortions, so that the welfare effects identified below are attributable to the supply disruption alone [21].

Facing shortages of the affected drug, pharmacists substitute toward the higher-price drug, expanding dispensing from Q_0 to Q_1. We deliberately represent this expansion as a quantity determined on the supply side rather than as a shift in patient demand: substitution decisions are made by pharmacists acting as supply-side intermediaries, not by patients expressing preferences, and the patient demand curve D remains unchanged throughout. Because the baseline allocation already exhausts all transactions whose value exceeds their cost, the entire expansion ΔQ lies beyond the socially efficient quantity. In this sense, the increase in higher-price drug dispensing is better understood as a form of supplier-induced demand, where the supplier, rather than the patient, determines the quantity transacted. In the standard supplier-induced demand literature, physician agency is typically modeled as a shift in the demand curve, reflecting physicians' ability to alter patient preferences (McGuire, 2000). Our setting differs in an important respect: patients are generally unaware that substitution has occurred at the point of dispensing, meaning that patient preferences are not merely manipulated but are effectively bypassed altogether. Because the quantity dispensed is determined unilaterally by pharmacists rather than by any demand-side mechanism, representing the quantity as chosen on the supply side, with the demand curve left intact, is the more appropriate representation of this institutional reality.

[21] In particular, with coinsurance, patients face only a fraction of P_H out of pocket, so patient self-selection alone would not place the baseline at E_1^*. The resulting moral-hazard wedge is a pre-existing distortion common to all insured drug purchases, which we deliberately set aside in order to isolate the effect of the disruption.

One may ask why the additional quantity is forthcoming at all, given that marginal cost exceeds the regulated price beyond Q_0. In the short run, the expansion is accommodated through existing inventories and strained production capacity, and manufacturers operating under shortage conditions face strong institutional pressure — including governmental requests to increase supply — to meet dispensing demand even when marginal units are unprofitable. More importantly for our purposes, the decomposition of the patient cost burden developed in Section A.2 depends only on the demand curve and the two regulated prices; it is unaffected by the shape of the marginal cost curve beyond Q_0, whose sole role is to pin down the efficiency benchmark E_1^*.

Because the higher-price drug is priced at $P_H > P_A$, patients who are switched from the affected drug to the higher-price drug face higher out-of-pocket spending. The increase in patient spending is represented by the shaded rectangle:

$$(P_H - P_A) \times \Delta Q,$$

where ΔQ denotes the increase in higher-price drug dispensing induced by the shortage. This rectangle measures the total increase in drug spending, which under Japan's universal public health insurance is divided between the public insurer and the patient: with a coinsurance rate of 30 percent, patients directly bear $0.3 \times (P_H - P_A) \times \Delta Q$ out of pocket, while the remaining share is absorbed by the public insurance system. The welfare decomposition developed in Section A.2 applies to the full rectangle and is invariant to this financing split; the split determines only the incidence of the burden between patients and the insurer. Importantly, this substitution occurs even when alternative, unaffected low-price drugs remain available in the market, suggesting that substitution reflects pharmacist discretion rather than a purely mechanical response to unavailability. As a result, the incidence of supply disruptions depends not only on the magnitude of the supply shock itself but also on the incentives facing pharmacists as intermediaries.

The graphical framework highlights two distinct channels through which supply disruptions reduce patient welfare. First, shortages in the affected-drug market reduce access to treatment. Second, substitution toward the more expensive higher-price drug increases out-of-pocket spending for patients who continue treatment. These channels apply to mutually exclusive groups of patients and thus involve no double counting: of the $Q_0 - Q'$ patients displaced from the affected drug, those switched to the higher-price drug (ΔQ) appear in the right panel and continue treatment at a higher cost, while the remainder forgo treatment and bear the consumer-surplus loss in the left panel. Accordingly, the shaded triangle should be read as the access loss of displaced patients who do not obtain a substitute. The second channel is not entirely a welfare loss, as these patients do obtain treatment. We will further investigate how the cost burden could be decomposed into the welfare components in Section A.2. To summarize, in regulated healthcare markets, both channels operate through non-price adjustment mechanisms rather than market-clearing price changes.

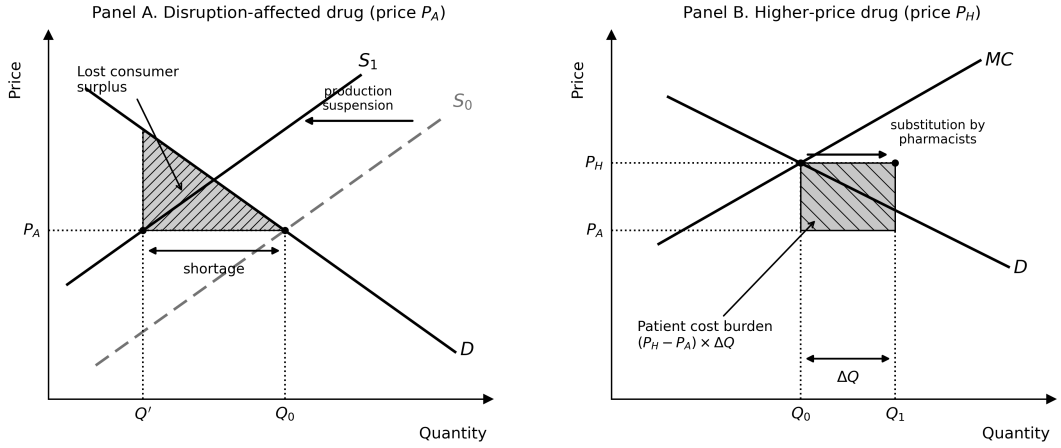


Figure A1: Supply Disruptions and Two Markets

A.2 Interpreting the Patient Cost Burden

Although our empirical analysis quantifies the increase in patients' out-of-pocket spending, this measure should not be interpreted mechanically as a complete welfare loss. A theoretical framework is therefore needed to interpret the estimated patient cost burden in welfare terms. Figure A2 provides an economic interpretation of the patient cost burden rectangle. The increase in spending is again given by $(P_H - P_A) \times \Delta Q$. A key feature of this framework is the role of pharmacists as supply-side intermediaries. In a frictionless market, dispensing would not extend beyond E_1^*, the quantity at which marginal willingness to pay equals marginal cost: units beyond this point reduce total surplus regardless of the prevailing price. Indeed, patients acting on their own preferences would accept units only up to the point where willingness to pay covers their out-of-pocket price; with insurance, this consent margin generally differs from E_1^*. Neither margin governs dispensing in our setting, however, because pharmacists have the discretion to substitute the higher-price drug without explicit patient consent. In our framework, the baseline allocation coincides with the efficiency benchmark — E_1^* is located at (Q_0, P_H) — so every unit of the substitution-induced expansion ΔQ is a unit whose production cost exceeds patients' willingness to pay. The location of Q_1 therefore reflects not patient demand but pharmacist-driven substitution, and the cost burden rectangle $(P_H - P_A) \times \Delta Q$ captures the full financial consequences of this intermediary behavior.

The demand curve divides this rectangle into two components corresponding to different economic interpretations. In region ②, below the demand curve, patient willingness to pay covers this portion of the payment: although these patients bear a higher out-of-pocket cost than under the affected drug's price, the payment is matched by the value of the treatment they receive. In region ①, above the demand curve, the payment exceeds patients' willingness to pay: patient valuation does not reach this portion of the payment, and in a market in which patients chose for themselves, these payments would never be

made. Region ① therefore represents spending that purchases no treatment value — a deadweight loss embedded in the patient cost burden, driven not by patient demand but by pharmacist discretion that bypasses patient preferences entirely.

Two remarks are in order. First, region ① captures the portion of the efficiency loss that is borne through patients' and the insurer's payments; any resource cost of supplying the marginal units in excess of P_H is borne by producers and lies outside the rectangle, so region ① is a conservative measure of the total efficiency loss from over-dispensing. Second, the welfare areas in Figures A1 and A2 are defined against different counterfactuals: the consumer-surplus triangle in the affected-drug market measures the incidence of the shock relative to the pre-shock allocation, whereas region ① measures the deadweight loss relative to the efficiency benchmark E_1^*. The two comparisons answer different questions — who bears the burden of the disruption, and how much social surplus is destroyed — and are therefore reported separately rather than summed.

In sum, the cost burden borne by patients and the insurer divides into a component that purchases genuine treatment value (②) and a component that constitutes pure deadweight loss (①). A central implication of this framework is that deadweight loss arises automatically whenever $\Delta Q > 0$: because the baseline allocation coincides with the efficiency benchmark E_1^*, any substitution-induced expansion necessarily pushes dispensing beyond the socially efficient quantity, so that region ① is strictly positive. The magnitude of this loss, however, is not mechanical — it depends on the slope of the demand curve. When demand is steep, willingness to pay falls quickly below P_H as dispensing expands, and a large share of the burden rectangle is deadweight loss; when demand is flat, marginal patients value the drug at nearly its price, and most of the burden purchases treatment value. The welfare cost of a given ΔQ is therefore an empirical question about the slope of demand, in addition to the extent to which pharmacist incentives drive substitution beyond E_1^*.

This decomposition highlights that increases in patient spending need not have a

uniform welfare interpretation. Higher spending may reflect payments for treatment that patients value, or spending that generates no value at all. More broadly, the framework illustrates how regulated intermediary markets can transform supply disruptions into patient financial burdens through non-price allocation mechanisms.

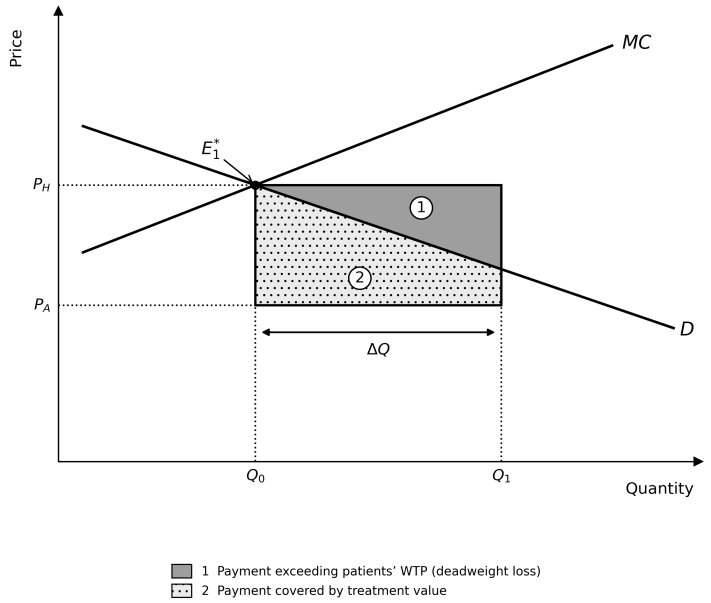


Figure A2: Decomposition of Cost Burden

B Pre-shock Price Tier of Suspended Generics

Figure 3 shows that the suspension increased dispensing of high-price generics, especially for hypertension. A potential concern is that this result may partly reflect differences in pre-shock price tiers rather than pharmacists' substitution decisions. If NP products were disproportionately located in the lower tier of the generic price distribution, then replacing suspended NP generics with remaining generics could mechanically raise high-price generic dispensing, particularly when comparable low-priced supplied generics were not available or not used. We therefore examine whether suspended NP generics were lower-priced than supplied generics within the same drug group and month before the suspension.

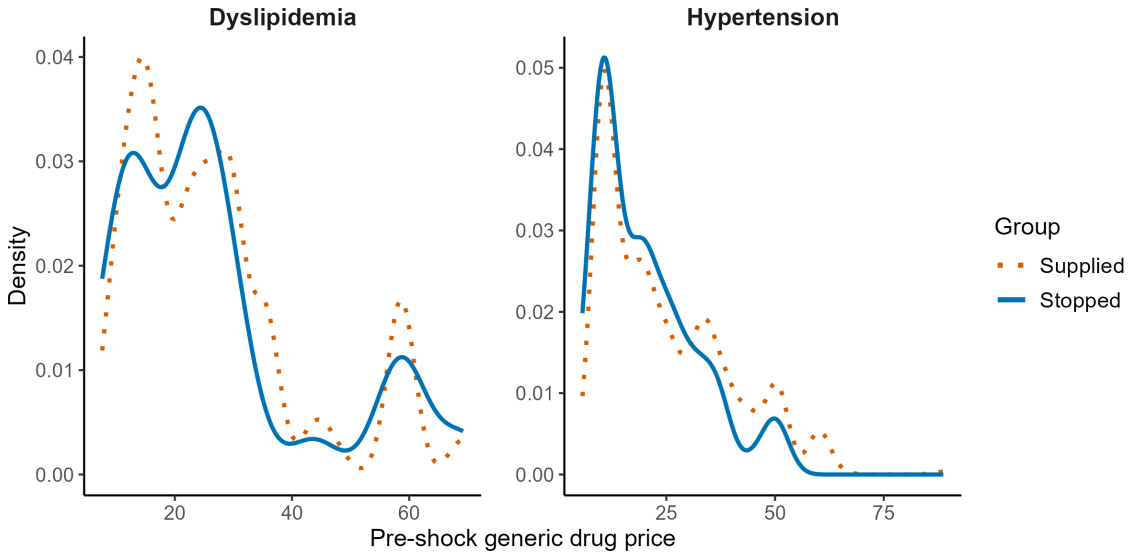
To assess this concern, we construct pre-suspension generic product-month data separately for each therapeutic class. The unit of observation is a generic product, identified by a YJ code, in month t. We collapse repeated claims for the same YJ code and month into one product-month observation. We restrict the sample to drug-group-by-month cells that contain at least one NP generic and at least one non-NP generic. We then estimate

$$Y_{jgt} = \beta D_{jgt} + \eta_{gt} + \varepsilon_{jgt}, \tag{5}$$

where j indexes YJ-code-level products, g indexes drug groups, t indexes months, and D_{jgt} equals one if product j is manufactured by NP. The term η_{gt} denotes drug-group-by-month fixed effects. We use two outcomes: (i) the log of the generic drug price and (ii) an indicator for whether the product is the highest-priced generic, defined as a generic whose price equals the maximum generic price within the same drug group and month. The first outcome measures whether suspended NP generics were lower-priced than supplied generics before the suspension. The second outcome is the same as the expensive-generic outcome in Figure 3; a negative estimate indicates that suspended NP generics were more

likely to be inexpensive generics before the suspension.

Figure A3: Pre-shock Generic Price Distributions



Notes: This figure plots the pre-shock distribution of generic drug prices separately for dyslipidemia and hypertension. The sample is restricted to generic product-month observations before the NP suspension and to drug-group-month cells that contain both suspended NP generics and supplied generics. Supplied generics are shown by the dotted red line, and stopped generics are shown by the solid blue line.

Appendix Figure A3 plots the raw pre-shock generic price distributions for suspended and supplied products. The distributions of suspended and supplied products overlap, in particular, for hypertension. Table A1 reports the results of corresponding regressions. For dyslipidemia, the estimated log price difference is small and statistically insignificant. For hypertension, suspended NP generics were about 8.4% lower-priced than supplied generics within the same drug group and month. Suspended NP generics were also 15.4 percentage points less likely to be high-price generics for hypertension.

These results suggest that the hypertension estimate in Figure 3 should be interpreted with caution. Suspended NP products for hypertension were less likely to be high-price generics before the suspension, implying that part of the post-suspension increase mechanically reflects the disappearance of relatively low-priced NP products. The size of this pre-shock price-tier imbalance is not negligible: the regression estimate is about half

Table A1: Pre-shock Price Differences

	Dyslipidemia		**Hypertension**	
Dependent Variable	Log(price)	High-price generic	Log(price)	High-price generic
NP Dummy	-0.038	-0.165	-0.087	-0.154
	(0.036)	(0.107)	(0.022)	(0.045)
Drug-group × month FE	×	×	×	×
Clustered SE	Drug group and month		Drug group and month	
Observations	2251	2251	7381	7381

Notes: This table reports product-month-level regressions using generic products observed before the NP suspension. The sample is restricted to drug-group-by-month cells that contain both NP and non-NP generics. NP Dummy equals one for NP generics and zero for non-NP generics. Highest-priced generic equals one if the product's regulated price equals the maximum generic price within the same drug group and month. All specifications include drug-group-by-month fixed effects. Standard errors are two-way clustered by drug group and month.

of the peak event-study response in Figure 3(b). Nevertheless, the remaining response is still comparable to the dyslipidemia response, supporting the interpretation that the suspension induced substitution toward higher-price generics in both therapeutic classes.

C Within-Generic Incentive Calculation

This section summarizes the construction of the within-generic incentive and savings reported in Table 3. The main calculation sets the unobserved procurement-cost differential, Δc, to zero. The required incentive therefore compensates only for the dispensing-profit loss from replacing a high-price generic with a lower-price alternative.

We infer product- and period-specific dispensing profits from the April 2022, April 2023, and April 2024 drug-price revisions. Let $P^{-}_{m,s}$ and $P^{+}_{m,s}$ denote the median regulated unit prices of product m immediately before and after revision s. Under the standard revision formula, the implied wholesale unit price is $\widehat{W}_{m,s} = P^{+}_{m,s} - 0.02P^{-}_{m,s}$, and the implied dispensing profit per unit is therefore $\widehat{\pi}_{m,s} = P^{-}_{m,s} - \widehat{W}_{m,s} = 1.02P^{-}_{m,s} - P^{+}_{m,s}$.

For the 2022 revision, the pre- and post-revision prices are measured over April 2021–March 2022 and April 2022–March 2023. The corresponding windows are April 2022–March 2023 and April 2023–March 2024 for the 2023 revision, and April 2023–March 2024 and April–September 2024 for the 2024 revision. If a product's regulated price did not change in the interim 2023 revision, we carry forward its preceding dispensing-profit estimate because the revision does not reveal a new wholesale price. We set negative implied profits to zero, cap them at the corresponding pre-revision regulated price, and do not separately model special repricing provisions.

Each estimate is assigned to the price regime from which it is inferred: the 2022 estimate to prescriptions from April 2021 through March 2022, the 2023 estimate to April 2022 through March 2023, and the 2024 estimate to April 2023 through March 2024. For observations before April 2021 that are required only for the Targeted DID calculation, we apply the 2022 dispensing-profit schedule as a common valuation benchmark. These earlier values are not interpreted as estimates of actual dispensing profits in those months.

For each drug group and month, the lower-price alternative is the generic with the lowest regulated price. If several products share the lowest price, we select the product

with the largest implied dispensing profit. We retain high-price generic prescriptions only when the drug group contains at least two generic price levels and dispensing profits are available for both products.

Let n index prescriptions, q_n denote the dispensed quantity, and g_n^H and g_n^L denote the high- and lower-price generics. Prescription-level dispensing profits, the raw profit gap, the required incentive, and the gross drug-spending reduction are

$$\widehat{\Pi}_n^H = q_n \widehat{\pi}_{g_n^H, t_n}, \quad \widehat{\Pi}_n^L = q_n \widehat{\pi}_{g_n^L, t_n}, \quad \widehat{G}_n = \widehat{\Pi}_n^H - \widehat{\Pi}_n^L, \quad \widehat{\tau}_n = \max\{\widehat{G}_n, 0\}, \quad \widehat{S}_n^T = q_n \left(P_n^H - P_n^L\right).$$

Thus, no incentive is required when the lower-price generic is already at least as profitable as the high-price generic. Patient and insurer gross savings are $\widehat{S}_n^P = 0.30\widehat{S}_n^T$ and $\widehat{S}_n^I = 0.70\widehat{S}_n^T$.

For any prescription-level quantity A_n, the Average columns report its simple mean among policy-relevant high-price generic prescriptions from April 2021 through March 2024. The Targeted columns measure the corresponding quantity per high-price generic prescription induced by the shortage. Let I_n indicate a policy-relevant high-price generic prescription, and let $\widehat{\beta}_d(Z)$ denote the treatment coefficient from the preferred DID specification for therapeutic class d when the outcome is Z_n. The Targeted estimate is

$$\widehat{A}_d^{\text{Targeted}} = \frac{\widehat{\beta}_d(IA)}{\widehat{\beta}_d(I)}.$$

For example, $\widehat{\tau}_d^{\text{Targeted}} = \widehat{\beta}_d(I\tau)/\widehat{\beta}_d(I)$ and $\widehat{S}_d^{T,\text{Targeted}} = \widehat{\beta}_d(IS^T)/\widehat{\beta}_d(I)$. The All–Average column pools observed prescriptions across the two therapeutic classes, whereas the All–Targeted column pools the class-specific DID numerators and denominators.

The incentive is treated as an insured dispensing fee. Patient and insurer contributions

are therefore $0.30\widehat{\tau}_d$ and $0.70\widehat{\tau}_d$, and the corresponding net savings are

$$\widehat{N}_d^P = \widehat{S}_d^P - 0.30\widehat{\tau}_d, \qquad \widehat{N}_d^I = \widehat{S}_d^I - 0.70\widehat{\tau}_d, \qquad \widehat{N}_d^T = \widehat{S}_d^T - \widehat{\tau}_d.$$

The incentive-to-gross-saving ratio is $\widehat{\tau}_d/\widehat{S}_d^T$. The reported break-even procurement-cost differential is $\widehat{\Delta c}_d = \widehat{S}_d^T - \widehat{\tau}_d$, treating each additional yen of procurement cost as requiring one additional yen of incentive. It is therefore an accounting threshold at which the net medical-spending reduction becomes zero, rather than an estimate of the unobserved procurement-cost differential.

Table A2: Implied Dispensing-Profit Components

	Dyslipidemia		Hypertension		All	
Measure	Average	Targeted	Average	Targeted	Average	Targeted
High-price generic implied profit	114	201	191	178	166	186
Lower-price generic implied profit	100	173	92	103	95	126
Raw profit gap ($\pi_{k,g^H} - \pi_{k,g^L}$)	13	28	99	76	71	60
Required incentive (τ)	24	28	101	75	77	60
Share with positive profit gap	86.7%	83.8%	96.9%	96.1%	93.6%	92.0%

Notes: All monetary quantities are in JPY per prescription. The raw profit gap is high-price generic profit minus lower-price generic profit. A positive gap indicates that replacing the high-price generic reduces the pharmacy's dispensing profit. The required incentive applies the positive-part operator to this gap with the procurement-cost differential set to zero. Average values are computed over observed policy-relevant high-price generic prescriptions from April 2021 through March 2024. Targeted values are ratios of the corresponding DID coefficients and should be interpreted as DID-implied quantities per shortage-induced prescription.

Table A2 reports the dispensing-profit components underlying these calculations. The raw profit gap is computed before truncation, whereas the required incentive applies the positive-part operator at the prescription level. In the Average columns, the mean required incentive therefore weakly exceeds the mean raw profit gap because negative gaps enter the latter but are replaced by zero in the former. The Targeted columns are ratios of DID coefficients rather than averages under a unique set of observed nonnegative prescription weights, so this inequality need not hold exactly. The final row reports the observed share

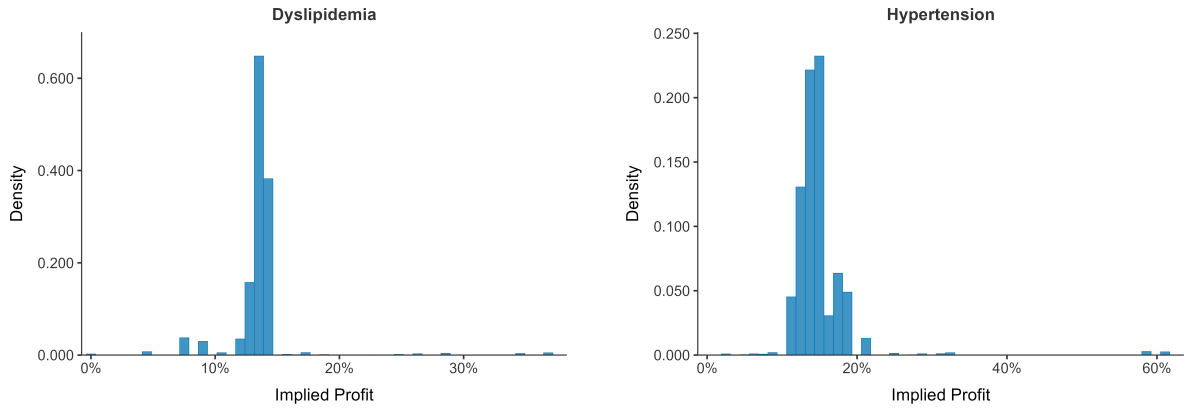
of prescriptions with a positive profit gap in the Average columns and the corresponding DID-implied ratio in the Targeted columns.

Appendix Figure A4 reports the distributions of the implied dispensing-profit rate among the policy-relevant high-price generic prescriptions used in the Average calculation. These distributions describe the underlying product-level profitability from which prescription-level profit gaps and required incentives are constructed.

Figure A4: Distribution of Implied Dispensing-Profit Rates

Panel (a): Dyslipidemia

Panel (b): Hypertension



Notes: The figure reports the claim-weighted distributions of the implied dispensing-profit rate among observed policy-relevant high-price generic prescriptions. The implied dispensing-profit rate is calculated at the product-by-revision level and assigned to prescriptions corresponding to each revision period. Products that are dispensed more frequently therefore receive greater weight in the distributions. Panel (a) reports the distribution for dyslipidemia, and Panel (b) reports the distribution for hypertension.

The implied profit rates are concentrated in the mid-teens in both therapeutic classes. The dyslipidemia distribution is relatively concentrated, whereas the hypertension distribution is more dispersed and has a longer upper tail. Differences in the required incentive across the two classes therefore depend not only on their central profit rates but also on the composition of products and dispensed quantities among policy-relevant prescriptions.

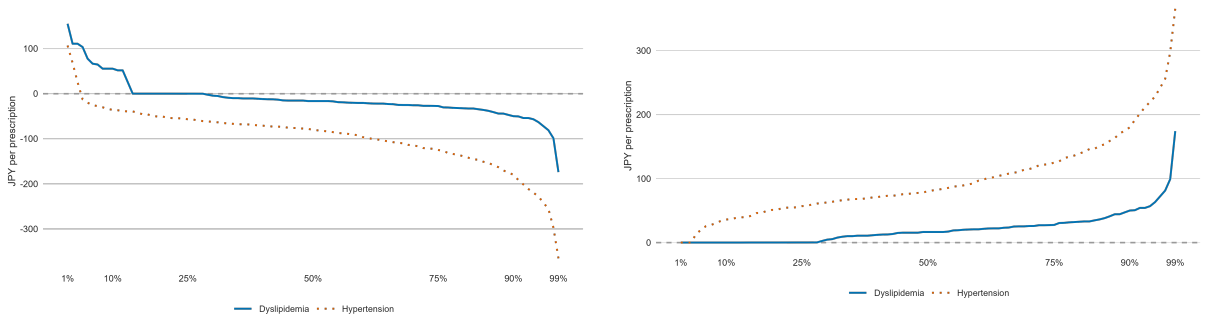
Appendix Figure A5 shows how differences between high- and lower-price generic profits translate into required incentives. Panel (a) plots lower-price generic profit minus high-price generic profit, which is the negative of the raw profit gap reported in Appendix

Table A2. Negative values in Panel (a) therefore correspond to prescriptions for which a positive incentive may be required.

Figure A5: Distributions of Pharmacy Profit Gaps and Required Incentives

Panel (a): Pharmacy profit-gap distribution

Panel (b): Required-incentive distribution



Notes: Panel (a) reports the percentile function of lower-price generic profit minus high-price generic profit, which is the negative of the raw profit gap reported in Table A2. Positive values indicate prescriptions for which the selected lower-price generic is more profitable. Panel (b) reports the percentile function of the minimum required incentive, defined as the positive part of high-price generic profit minus lower-price generic profit when the unobserved procurement-cost differential is set to zero. The distributions are constructed from observed policy-relevant high-price generic prescriptions between April 2021 and March 2024. Monetary values are measured in JPY per prescription. The solid blue line denotes dyslipidemia, and the dotted orange line denotes hypertension.

The lower-price generic is less profitable for most policy-relevant prescriptions in both therapeutic classes, but this difference is more prevalent and larger in hypertension. Applying the positive-part operator produces a substantial mass at zero for dyslipidemia, while positive incentives arise over a larger portion of the hypertension distribution and increase more sharply in its upper tail. These distributional differences account for the larger Average required incentive for hypertension reported in Table 3.

D Additional Figures and Tables

Table A3: Descriptive statistics of pharmacies

	Mean	Std. Dev.	Min	Max
Number of pharmacies	22,469	1,926	18,635	25,299
Share of add-on categories				
No	0.176	0.031	0.129	0.242
Low	0.201	0.033	0.145	0.244
Middle	0.371	0.005	0.358	0.381
High	0.252	0.068	0.137	0.368
Chain store	0.325	0.013	0.310	0.352

Notes: The table reports summary statistics across months from June 2020 through September 2024. The number of pharmacies is the monthly count of pharmacies appearing in the claims data. The add-on categories and chain-store status are reported as monthly shares of observed pharmacies. The sample includes both therapeutic classes and products manufactured by all firms.

Table A4: Excluded drug groups

Therapeutic class	YJ code and drug group	Reason for exclusion
Dyslipidemia	2189018F1: Ezetimibe tablets 10 mg	A generic entry occurred during the post-patent-expiration analysis period, changing the product choice set.
Dyslipidemia	2189015F1: Atorvastatin tablets 5 mg 2189015F2: Atorvastatin tablets 10 mg 2189015F5: Atorvastatin tablets 20 mg 2189011F3: Simvastatin tablets 20 mg	These drug groups were used only by one of the exposure groups in the estimation sample.
Dyslipidemia	2189017F1: Rosuvastatin tablets 2.5 mg 2189017F3: Rosuvastatin orally disintegrating tablets 2.5 mg	The control group experienced new product entry unrelated to the NP suspension.
Hypertension	2149044F4: Olmesartan tablets 40 mg 2149044F5: Olmesartan orally disintegrating tablets 10 mg 2149044F6: Olmesartan orally disintegrating tablets 20 mg	The product line had been discontinued as of June 2020, changing the product choice set for reasons unrelated to the NP suspension.
Hypertension	2144006F2: Lisinopril hydrate tablets 10 mg	Supply was temporarily suspended in October 2020 because a pre-shipment manufacturing lot failed to meet the dissolution specification.
Hypertension	2149118F1: Irbesartan and amlodipine combination tablets LD	The same product was also sold by other firms, placing it outside the treatment definition based on Nichi-Iko products.

Notes: This table lists drug groups excluded from the estimation sample. The exclusions remove drug groups whose product choice sets or pricing environments changed for reasons unrelated to the NP production suspension. The excluded groups include drug groups with generic entry during the analysis period, drug groups used only by one exposure group, drug groups with abrupt price or cost discontinuities, drug groups affected by product-line discontinuation or temporary supply suspension, and drug groups outside the treatment definition.

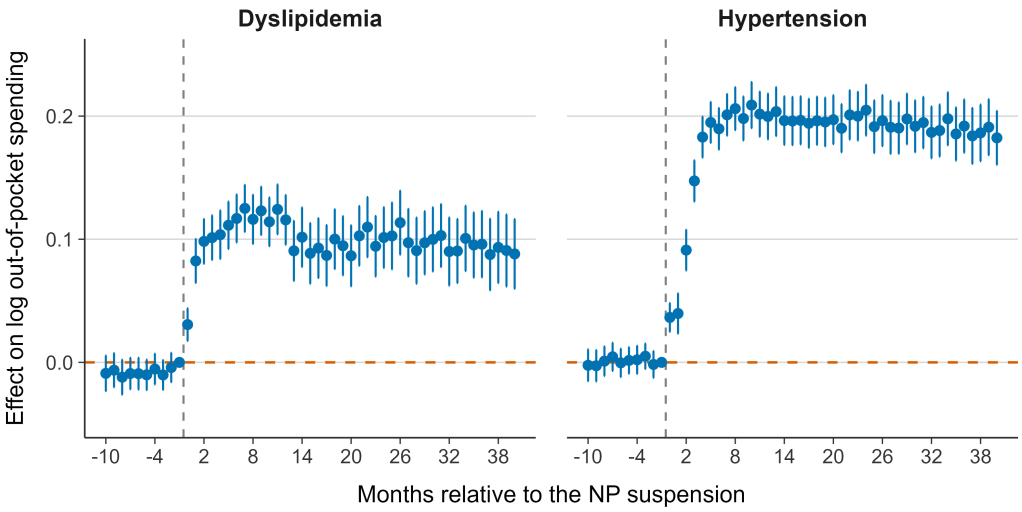
Table A5: Difference in the impact on the patient's copayment by pharmacy type

	Dyslipidemia		**Hypertension**	
	(1)	**(2)**	**(3)**	**(4)**
δ No Add-on	0.034	0.006	-0.005	-0.018
	(0.009)	(0.009)	(0.008)	(0.007)
Individual FE	×	×	×	×
Month FE	×	×	×	×
Pharmacy FE	×	×	×	×
Drug-group FE	×	×	×	×
FY × Pharmacy FE	×	×	×	×
Month × Drug-group FE		×		×
Observations	1,379,895	1,379,895	2,091,311	2,091,275
N patients	64,588	64,588	93,375	93,375

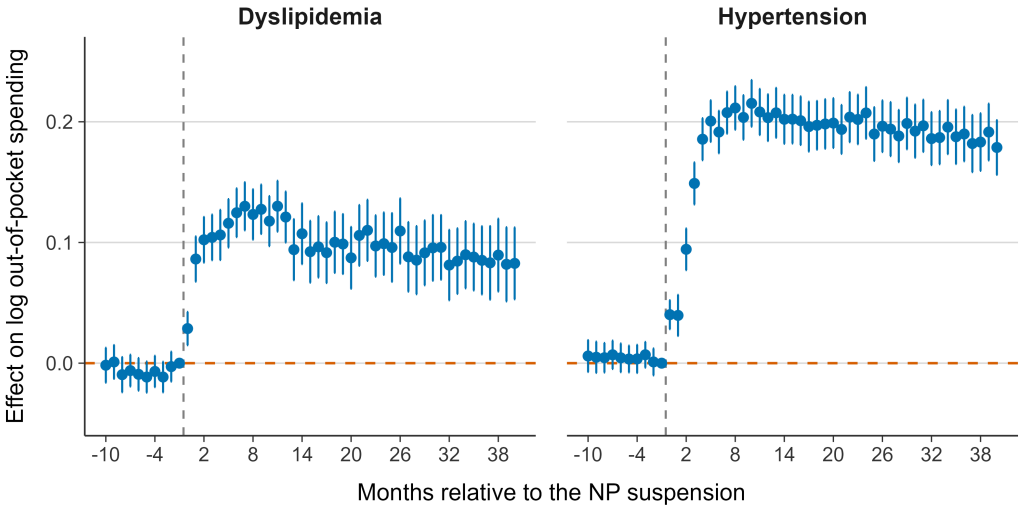
Notes: The table reports the coefficient of the interaction term of δ and a dummy variable to indicate the level of supplementary fees to promote generic drugs that were included in the post-period DID specification, and their standard errors. In this analysis, a dummy variable at the pharmacy level is defined, depending on whether the pharmacy receives a generic supplementary fee. The reference group is the pharmacy that receives the add-on. The outcome variable Y_{ijkt} denotes the patient's copayment of the dispensed drug. Odd-numbered columns control for individual and month fixed effects; even-numbered columns additionally include pharmacy fixed effects and their interaction with fiscal-year dummies.

Figure A6: The effect on out-of-pocket spending

Panel (a): generic share threshold 0.7



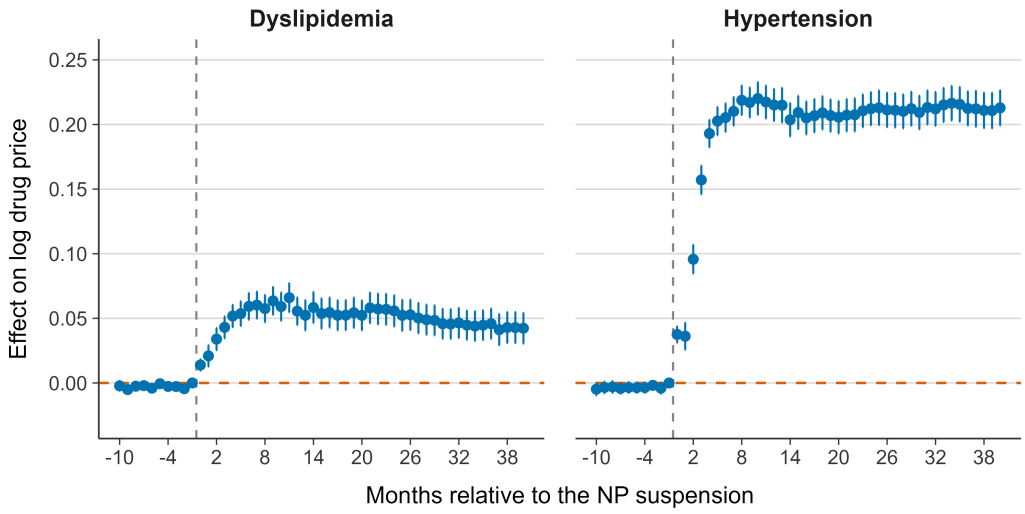
Panel (b): generic share threshold 0.9



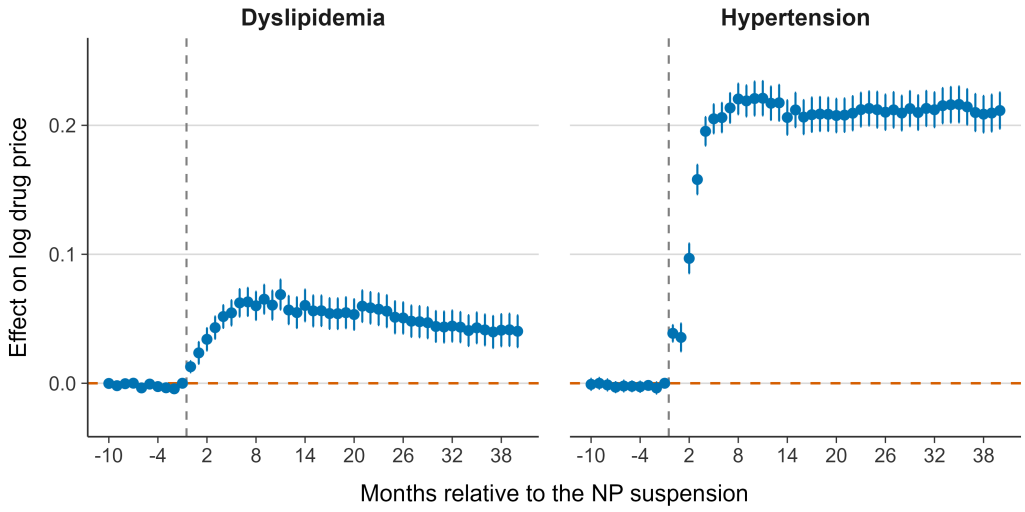
Notes: The figure reports event-study estimates and 95% confidence intervals. The outcome is the log of patient out-of-pocket spending per prescription. All specifications include patient-by-drug-group fixed effects, month-by-drug-group fixed effects, and fiscal-year-by-pharmacy fixed effects. Standard errors are two-way clustered at the patient and pharmacy levels. The vertical dashed line indicates the NP production suspension, and the omitted event month is one month before the suspension. Panel (a) defines treatment exposure using a pre-suspension generic-use threshold of 0.7, and Panel (b) uses a threshold of 0.9.

Figure A7: The effect on drug prices

Panel (a): generic share threshold 0.7



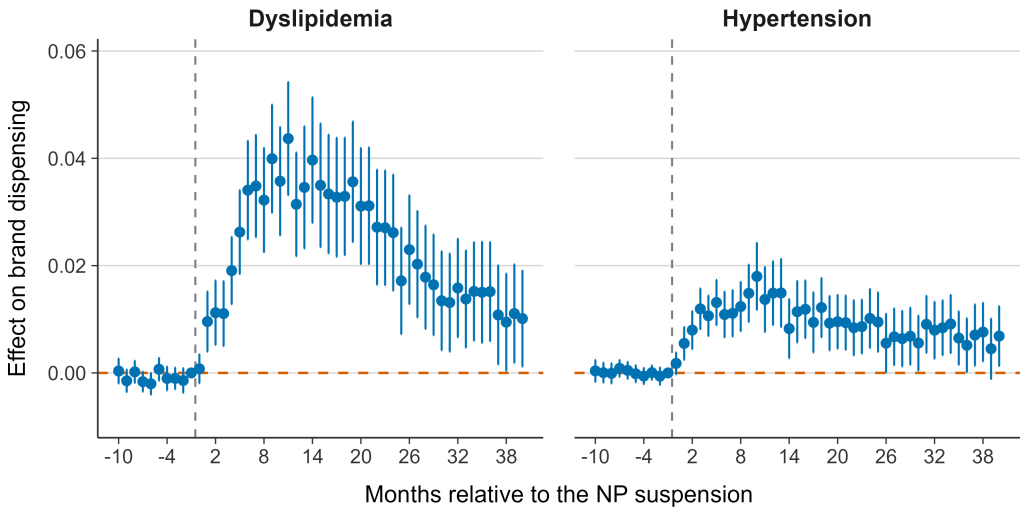
Panel (b): generic share threshold 0.9



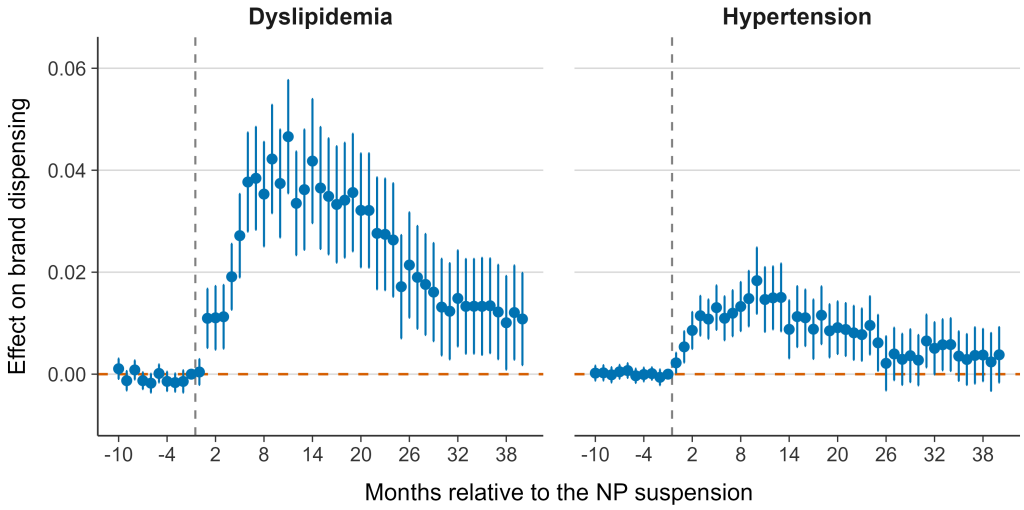
Notes: The figure reports event-study estimates and 95% confidence intervals. The outcome is the log of the regulated drug price per dose. All specifications include patient-by-drug-group fixed effects, month-by-drug-group fixed effects, and fiscal-year-by-pharmacy fixed effects. Standard errors are two-way clustered at the patient and pharmacy levels. The vertical dashed line indicates the NP production suspension, and the omitted event month is one month before the suspension. Panel (a) defines treatment exposure using a pre-suspension generic-use threshold of 0.7, and Panel (b) uses a threshold of 0.9.

Figure A8: The effect on brand-name drug substitution

Panel (a): generic share threshold 0.7



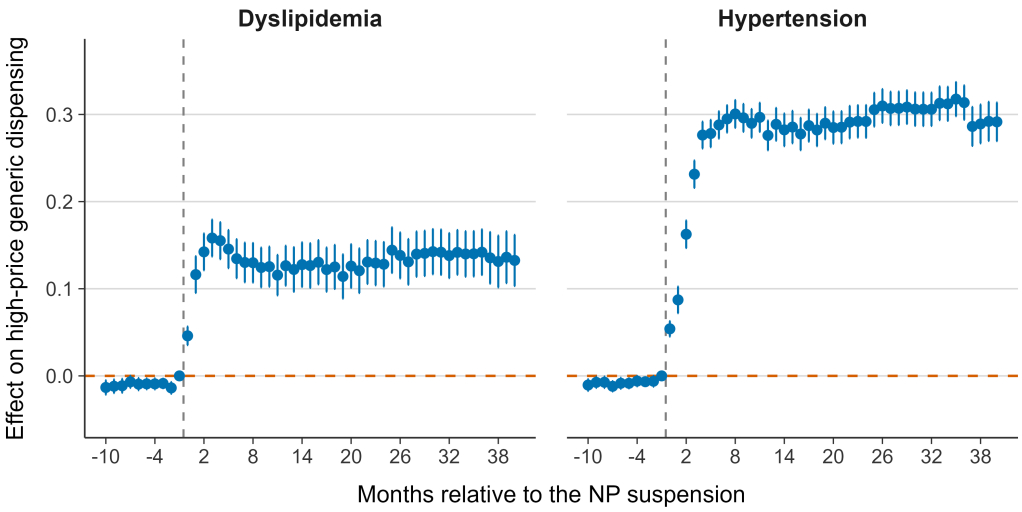
Panel (b): generic share threshold 0.9



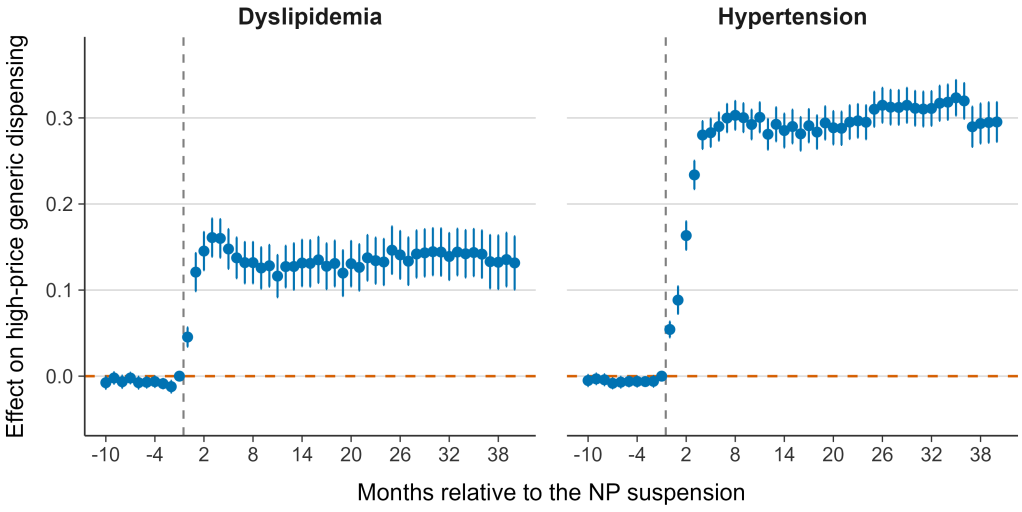
Notes: The figure reports event-study estimates and 95% confidence intervals. The outcome is an indicator equal to one if the dispensed product is a brand-name drug. All specifications include patient-by-drug-group fixed effects, month-by-drug-group fixed effects, and fiscal-year-by-pharmacy fixed effects. Standard errors are two-way clustered at the patient and pharmacy levels. The vertical dashed line indicates the NP production suspension, and the omitted event month is one month before the suspension. Panel (a) defines treatment exposure using a pre-suspension generic-use threshold of 0.7, and Panel (b) uses a threshold of 0.9.

Figure A9: The effect on high-price generic substitution

Panel (a): generic share threshold 0.7

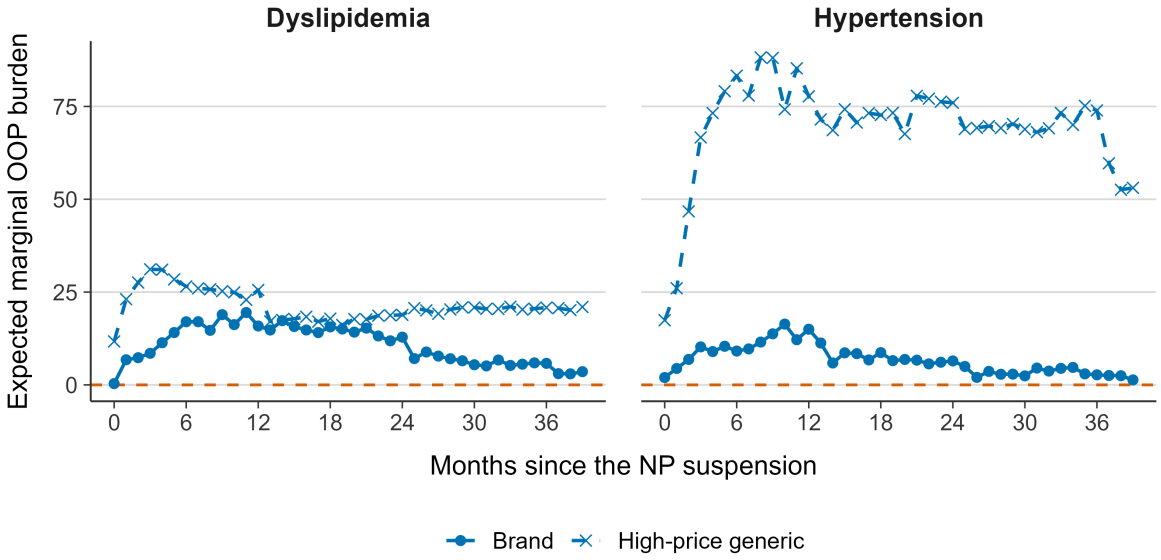


Panel (b): generic share threshold 0.9



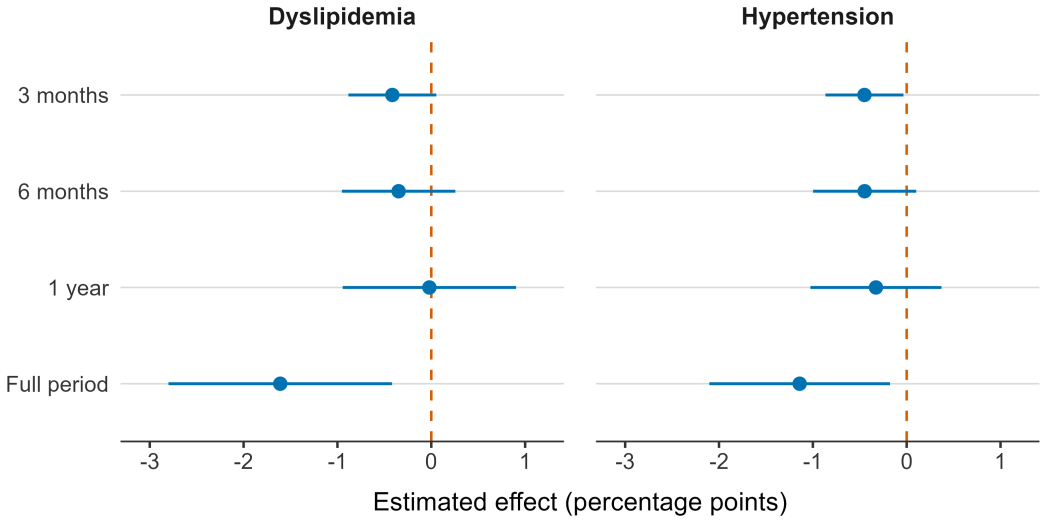
Notes: The figure reports event-study estimates and 95% confidence intervals. The outcome is an indicator equal to one if the dispensed product is the highest-priced generic within its drug group and period. All specifications include patient-by-drug-group fixed effects, month-by-drug-group fixed effects, and fiscal-year-by-pharmacy fixed effects. Standard errors are two-way clustered at the patient and pharmacy levels. The vertical dashed line indicates the NP production suspension, and the omitted event month is one month before the suspension. Panel (a) defines treatment exposure using a pre-suspension generic-use threshold of 0.7, and Panel (b) uses a threshold of 0.9.

Figure A10: Effect on Marginal Patient Costs



Notes: This figure reports the marginal contribution of substitution to branded drugs and to the highest-priced generic drugs to the increase in patients' out-of-pocket spending in each event month. For each event month and drug group, we calculate the average out-of-pocket price gap relative to the lowest-priced generic using prescription quantities as weights. The marginal burden is obtained by multiplying these price gaps by the corresponding event-study estimates of branded and highest-priced generic dispensing shown in Figure 3.

Figure A11: Difference in the patient's pharmacy change



Notes: The figure reports estimates from $Y_i = \alpha + \beta D_i + X_i'\gamma + \varepsilon_i$, where Y_i is an indicator equal to one if patient i uses at least one pharmacy during the corresponding post-suspension period that was not used before the suspension. The indicator D_i equals one if patient i used NP products before the suspension within the corresponding therapeutic class. The covariate vector X_i includes the patient's pre-suspension age, sex, and number of distinct pharmacies used during the pre-suspension period. Age is measured at the patient's last observed dispensing before the suspension. Standard errors are clustered at the pre-suspension pharmacy level. Horizontal bars denote 95% confidence intervals.