Cabinet Office

**ESRI Discussion Paper Series No.396**

# Impact of Hospital Decision-making on Drug Markets: The Case of Biosimilars

Kyogo Kanazawa

July 2024

Economic and Social Research Institute
Cabinet Office
Tokyo, Japan

# Impact of Hospital Decision-making on Drug Markets: The Case of Biosimilars*

Kyogo Kanazawa†

July 11, 2024

**Abstract**

In the healthcare industry, governments regulate prices for several reasons. To appropriately regulate prices for policy objectives, it is necessary to understand the response of medical demand to price changes. This study analyzed the Japanese pharmaceutical market for Filgrastim, an expensive biologic drug, and its generics (biosimilars) using a demand model that explicitly incorporates the introduction of generics by hospitals, using data from 184,954 hospitalized patients. We analyzed the impact of both original and generic drug prices on the introduction of generic drugs in hospitals using the hazard model as a first step and estimated a discrete choice model for each patient's drug choice as a second step; however, the choice set for each patient was restricted to the drugs that were available at their hospital at that time. Then, using the results of both steps, we conducted a counterfactual simulation and showed that a 10% reduction in the price of generic drugs would increase its market share by 12.04% in the most recent year and that this spreading effect mostly comes from the hospitals' decision to introduce generics.

**Keywords:** Biologic; Biosimilar; Generic Drug; Hospital; Pharmaceutical; Japan.
**JEL Codes:** I11; K23; L65; O34.

*I am particularly grateful to my adviser Hidehiko Ichimura for his invaluable guidance and continuous support. I also thank Professors Toshiaki Iizuka, Haruko Noguchi, Naoki Wakamori, and Yasutora Watanabe, who were members of my dissertation review committee, and participants of the seminars at Keio University, Otaru University of Commerce, JEA, Happy-Hour Seminar, Yokohama City University, AWEHE, Yokohama National University, and ESRI for their constructive comments. The Diagnosis Procedure Combination (DPC) database used in this paper is provided by Hideo Yasunaga and Kiyohide Fushimi, and financially supported by Grant-in-aid from the Ministry of Health, Labour, and Welfare, Japan (grant number: H30-Policy-Designated-004). I also thank Hiroki Matsui and Yusuke Sasabuchi for their advice, particularly regarding the DPC database.

†kanazawa-kyogo-kb@ynu.ac.jp, Faculty of International Social Sciences, Yokohama National University

# 1 Introduction

In recent years, the increase in medical costs due to the aging population and the sophistication of medical technology has become an issue in many developed countries. Therefore, controlling medical costs while taking into consideration the welfare of patients, profits of medical institutions, and profits and development incentives of pharmaceutical companies and medical device manufacturers is important. To address this problem, governments in many countries including Japan, the focus of this study, regulate prices of medical practices and pharmaceuticals. For governments to appropriately regulate prices for policy objectives, it is first necessary to properly understand the response of demand to price changes.

In demand analysis of the pharmaceutical market, a common model is that a patient or doctor chooses a drug from the drugs available in the market at that time. However, regarding prescriptions in hospitals, the board of directors of a hospital, for example, decides in advance the drugs to purchase, and the prescriptions for patients are selected from among those drugs. Therefore, using a demand model that explicitly incorporates the introduction of generics by hospitals, this study analyzes the Japanese market for Filgrastim, an expensive biologic drug mainly used for neutropenia in patients with cancer, and its generics (biosimilars, BS), using data from 184,954 hospitalized patients in Japan.

First, we analyzed the impact of original and generic drug prices on the introduction of generic drugs in hospitals using a hazard model with each hospital as the data unit. The results showed that a higher price for the original drug accelerated the introduction of generics, whereas a higher price for generics decelerated the introduction of generics.

Second, we estimated a discrete choice model for each patient's drug choice. The estimation method was based on the method of Dunn (2012), including its instrumental variables; however, the choice set for each patient was restricted to the drugs that were available at their hospital at that time. The results showed that drug prices did not have a significant effect on patients' drug choices.

Finally, using the results of the first and second models of the estimation, we conducted a counterfactual simulation to see the effect of changes in the price of generics on their market share. The results showed that a 10% reduction in the price of generic drugs increased their market share by 12.04% in the most recent year. Additionally, we conducted several simulations to decompose the factors of this result and showed that this spreading effect mostly comes from hospitals' decisions to introduce generics.

This study contributes to existing literature in two ways. First, this study is based on the discrete choice model of Berry (1994) and Berry et al. (1995), a standard method in the empirical industrial organization field, and its applications to pharmaceuticals (Cleanthous

(2002); Iizuka (2007); Dunn (2012); Duso et al. (2014); Kaiser et al. (2014); Dubois and Lasio (2018)). This study contributes to the literature by estimating the choice set of each agent in part. While the hazard model in this field is mainly used to analyze the timing of new drugs entering the market (Danzon et al. (2005); Cockburn et al. (2016)), we used it to analyze the timing of the introduction of new generic drugs by hospitals.

Second, this study contributes to the literature on hospital behavior. Existing studies on hospital-level behavior have analyzed price negotiations with suppliers (Grennan and Swanson (2020)), changes in behavior due to changes in revenue policies (Duggan (2000); Dafny (2005)), and responses to economic recession (Dranove et al. (2017)). However, to the best of our knowledge, no existing studies have analyzed drug adoption behavior at the hospital level, and this is the first study to do so.

The reminder of this paper is organized as follows. Section 2 describes the industry background of the market. Section 3 describes the hospitalization data and summarizes the statistics. Section 4 explains and conducts hazard model regressions of generics introduction by hospitals. Section 5 describes and conducts the discrete drug choice model of patients. Section 6 conducts simulations using the estimation results of the previous two sections and compares the simulation results with the actual data and other simulation results using the patients' drug choice model only. Section 7 presents the conclusion of the study.

# 2 Industry Backgrounds

## 2.1 Focused Drug

In this study, we focus on the biologic drug Filgrastim, its generic drugs (biosimilars), and other competitive pharmaceuticals.

Filgrastim is a biologic drug primarily used in patients with neutropenia for chemotherapy. In Japan, Filgrastim was approved in 1991 and three types of Filgrastim biosimilars were approved between 2013 and 2014. It was the first biologic drug to have more than one biosimilar in Japan. Other competing drugs for neutropenia include Nartograstim, Lenograstim, Mirimostim, and Pegfilgrastim. In this study, 31 products of different sizes and dosage forms were analyzed.

Biologic drugs are pharmaceuticals that "are derived from proteins and other substances produced by living organisms, such as mammalian cells, viruses and bacteria" (International Federation of Pharmaceutical Manufacturers & Associations (2012)). In contrast to most traditional, small-molecule drugs, biologic drugs require more sophisticated technologies for manufacturing and quality control and therefore tend to be priced higher. In 2016, biologic

drugs occupied about a quarter of global pharmaceutical market sales, and it seems that this proportion will continue to increase (Evaluate Pharma (2017)).

Biosimilars are generic drugs for biologic drugs. To show high similarity to an already approved biologic, a biosimilar product needs to meet more rigorous standards of safety and efficacy, accompanied by more clinical tests, than a generic product of a small-molecule drug. In Japan, only 11 biosimilars were approved as of 2017 although over 130 biologic drugs existed. As patents for biologic drugs will expire in the future, it seems that the entries of biosimilars will increase.

## 2.2 Drug Price in Japan

In Japan, the retail prices of pharmaceuticals are open to everyone and deterministically determined by the government based on their ex-retail price and ex-average wholesale price. Therefore, it is possible to recover the average wholesale price from that formula. During the sample period, price revisions were made every two years (2012, 2014, 2016, and 2018) in April. As in Iizuka (2007) and Iizuka (2012), we can restore the unobserved average wholesale prices from observed retail prices.

Specifically, if the retail price is revised in year $t+2$, the next retail price of the drug $j$ is determined as follows

$$P^R_{j,t+2} = P^W_{jt} \times (1 + r_{\text{tax},t+2}) + 0.02 \times P^R_{jt}$$

where $P^R$ indicates the retail price (including tax), $P^W$ indicates the average wholesale price (excluding tax), and $r_{\text{tax}}$ indicates the consumption tax rate (5% until March 2014 and 8% since April 2014). Therefore, the average wholesale price, including the tax on drug $j$ in year $t$, is given by

$$P^W_{jt} \times (1 + r_{\text{tax},t}) = (P^R_{j,t+2} - 0.02 \times P^R_{jt}) \times (1 + r_{\text{tax},t})/(1 + r_{\text{tax},t+2}). \tag{1}$$

Table 1 shows the retail prices of the focused drugs in the sample period. For example, the retail price of Filgrastim (original) syringe 75$\mu$g was 10,055 JPY in FY2012-13 and 9,481 JPY in FY2014-15, thus, the average wholesale price including tax for FY2012-13 is obtained by $(9481 - 0.02 * 10055) * 1.05/1.08 = 9022.125$ JPY. Figure 1 shows the graphs of price transitions of Filgrastim (original) syringe 75$\mu$g. The transition of the wholesale price is similar to that of the retail price, and the retailers' markups (retail price - wholesale price) did not move significantly.[1]

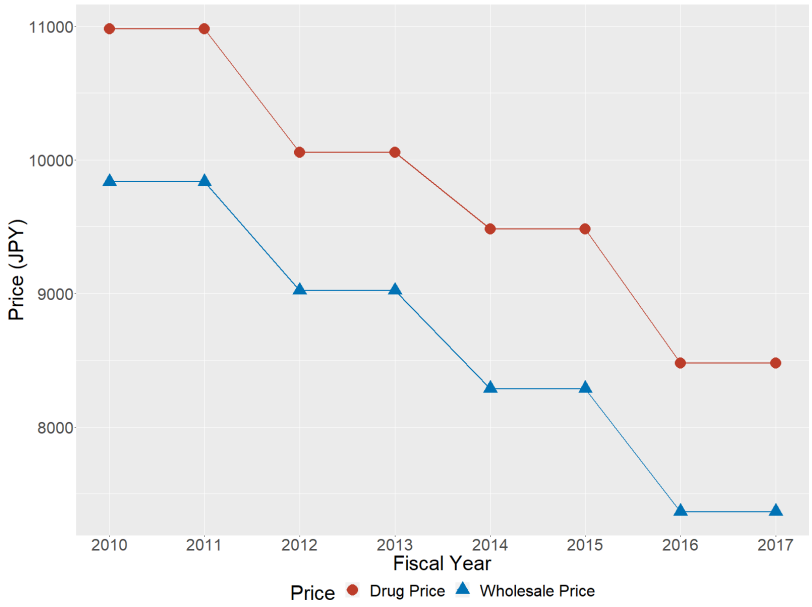
[1]Since a different system is applied, the average wholesale price of Pegfilgrastim cannot be calculated

Table 1: Drug Prices

| Drug \FY | 2010-11 | 2012-13 | 2014-15 | 2016-17 | 2018-19 |
|---|---|---|---|---|---|
| Filgrastim (original) Syringe 75$\mu$g | 10,981 | 10,055 | 9,481 | 8,477 | 7,536 |
| Filgrastim (original) Syringe 150$\mu$g | 21,811 | 20,048 | 18,936 | 16,961 | 15,091 |
| Filgrastim (original) Syringe 300$\mu$g | 27,325 | 24,926 | 23,542 | 20,969 | 18,580 |
| Filgrastim (original) Injection 75$\mu$g | 10,908 | 10,055 | 9,481 | 8,635 | 7,845 |
| Filgrastim (original) Injection 150$\mu$g | 22,270 | 20,048 | 18,900 | 17,624 | 15,910 |
| Filgrastim (original) Injection 300$\mu$g | 27,027 | 24,781 | 23,475 | 21,117 | 19,168 |
| Filgrastim (F) 75$\mu$g | - | 6,882 | 6,143 | 5,071 | 4,069 |
| Filgrastim (F) 150$\mu$g | - | 10,871 | 9,987 | 8,238 | 6,497 |
| Filgrastim (F) 300$\mu$g | - | 17,179 | 15,093 | 12,525 | 10,163 |
| Filgrastim (Mochida) 75$\mu$g | - | 6,882 | 6,143 | 5,071 | 4,069 |
| Filgrastim (Mochida) 150$\mu$g | - | 10,871 | 9,987 | 8,238 | 6,497 |
| Filgrastim (Mochida) 300$\mu$g | - | 17,179 | 15,093 | 12,525 | 10,163 |
| Filgrastim (NK) 75$\mu$g | - | 6,882 | 6,143 | 5,071 | 4,069 |
| Filgrastim (NK) 150$\mu$g | - | 10,871 | 9,987 | 8,238 | 6,497 |
| Filgrastim (NK) 300$\mu$g | - | 17,179 | 15,093 | 12,525 | 10,163 |
| Filgrastim (Teva) 75$\mu$g | - | 6,882 | 6,143 | 5,071 | 4,069 |
| Filgrastim (Teva) 150$\mu$g | - | 10,871 | 9,987 | 8,238 | 6,497 |
| Filgrastim (Teva) 300$\mu$g | - | 17,179 | 15,093 | 12,525 | 10,163 |
| Filgrastim (Sandoz) 75$\mu$g | - | - | 6,143 | 3,971 | 2,657 |
| Filgrastim (Sandoz) 150$\mu$g | - | - | 9,987 | 8,238 | 6,497 |
| Filgrastim (Sandoz) 300$\mu$g | - | - | 15,093 | 9,757 | 5,526 |
| Lenograstim 50$\mu$g | 6,155 | 5,685 | 5,430 | 4,749 | 4,079 |
| Lenograstim 100$\mu$g | 11,339 | 10,445 | 9,907 | 8,542 | 7,388 |
| Lenograstim 250$\mu$g | 28,144 | 25,881 | 24,566 | 21,098 | 18,197 |
| Mirimostim 400 | 13,370 | 13,287 | - | - | - |
| Mirimostim 800 | 26,095 | 23,897 | 23,293 | 20,434 | 18,154 |
| Nartograstim 25$\mu$g | 4,732 | 4,659 | 4,421 | 3,937 | 3,531 |
| Nartograstim 50$\mu$g | 9,564 | 8,712 | 8,225 | 7,252 | 6,432 |
| Nartograstim 100$\mu$g | 19,064 | 17,362 | 16,256 | 14,396 | 13,044 |
| Nartograstim 250$\mu$g | 29,284 | 28,125 | 26,348 | 23,800 | 21,754 |
| Pegfilgrastim 3600$\mu$g | - | - | 106,660 | 106,660 | 106,660 |

Notes: The data source is the Ministry of Health, Labour, and Welfare, Japan. These prices are in JPY (100 JPY was roughly 1 USD at that time). "Filgrastim (original)" refers to the original filgrastim product produced by Kyowa Hakko Kirin, and other "Filgrastim (...)" refers to the biosimilar products of filgrastim, with the names of the producing companies in parentheses.

Figure 1: Drug Price and Wholesale Price: Filgrastim (original) Syringe 75$\mu$g




Notes: The data source is the Ministry of Health, Labour, and Welfare, Japan. These prices are in JPY (100 JPY was roughly 1 USD at that time). "Filgrastim (original)" refers to the original filgrastim product produced by Kyowa Hakko Kirin. Wholesale prices were calculated using Equation (1).

However, we could not restore the wholesaler price for each biosimilar product due to the different pricing systems for biosimilars. The first biosimilar to be approved is priced at 70% of the original price, and subsequent biosimilars are priced at the same level as the lowest-priced biosimilar at the time of approval. In addition, when the drug price is revised, the new price calculated based on the weighted average of all biosimilars is applied to all biosimilars. Therefore, the wholesale price for each biosimilar company could not be calculated, and only the average wholesale price of all biosimilars was calculated.

Next, we clarify the reasons for using wholesale prices in our analysis.

## 2.3 DPC/PDPS

We focused on inpatients at Japanese hospitals participating in the Diagnosis Procedure Combination/Per-Diem Payment System (DPC/PDPS), a comprehensive evaluation system of medical fees for acute inpatient care, in this study. Approximately 1,000 hospitals in Japan participate in the DPC/PDPS, and all of them cover approximately 50% of all inpatients in

using this method. Therefore, the average wholesale price of Pegfilgrastim was calculated by extrapolating the markup rate of Filgrastim (original) syringe 75$\mu$g, the most popular product of original Filgrastim and produced by the same company Kyowa Hakko Kirin.

acute care beds in Japan.

Notably, the payment of medical fees for acute inpatient care in the DPC/PDPS is the sum of the comprehensive evaluation portion and piece-rate evaluation portions in participating hospitals. The comprehensive evaluation portion is calculated based on the number of points per day set for each patient's DPC and the number of days hospitalized. Hence, the actual treatment given to the patient is not directly related to the medical fee. As prescription drugs are included in the comprehensive evaluation portion, the retail price of drugs in hospitals participating in this system does not affect either the hospital's revenue or the patient's cost; only the wholesale price affects the hospital's cost. Therefore, in the subsequent analyses, the model was formulated by assuming that only wholesale prices matter.

# 3 Data

In this study, we used the Diagnosis Procedure Combination (DPC) database, which consists of inpatients at Japanese hospitals participating in the DPC/PDPS. The data consisted of hospitalization episode units and data such as hospital ID, patient ID, admission date, discharge date, patient's sex, birthday, height, weight, disease codes (ICD-10), prescription drugs, and prescription date. One disadvantage is that the patient ID is unique only to the hospital in which the patient is admitted. If the same patient was admitted to the same hospital multiple times during the data period, it was identified as the same person. However, if the patient was admitted to a different hospital, they could not be identified as the same person.

The main target of the drugs to be analyzed is neutropenia, but there exists a large number of patients who use one or more of these drugs without a corresponding disease code (D70 in ICD-10). In addition, as this drug is used by patients with a wide variety of diseases, and the demand for this drug can vary greatly depending on the disease, we limited the sample of patients with representative diseases in terms of consumption as much as possible. Therefore, we used the following procedure to extract patient and hospitalization data for the analysis:

First, we obtained all hospitalization records in which the ICD-10 chapter of the main disease was C00-D48 (neoplasms) or D50-D89 (diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism). Subsequently, we summed the total amount of focused drugs prescribed for each ICD-10 code in the middle category of the main disease in these hospitalization records. Consequently, the hospitalization records of patients with a main disease of C81-C96 (malignant neoplasms, stated or presumed to be primary, of lymphoid, haematopoietic and related tissue) accounted for the majority of

prescriptions (55.01%); therefore we chose these records as our analyzed sample.

Second, from the hospitalization data with the codes C81-C96 as the main disease, we determined the sample hospitalization and reference date for each patient using the following procedure. For patients who used the focused drug at least once, the first hospitalization in which the drug was used was selected as the sample hospitalization, and the date on which the drug was first prescribed during hospitalization was used as the reference date. For patients who had never used the focused drugs, the first hospitalization within the sample period was considered as the sample hospitalization. The median number of days until the first prescription for each number of days of hospitalization was calculated from the data of the patients for whom the focused drug was prescribed, and the median days corresponding to the number of days of hospitalization for that hospitalization after the date of admission became the reference date for that patient.

Based on the data of sample hospitalizations and reference dates, we excluded patients whose hospitalization started before June 2010, missing or abnormal values for height, weight, or birthday, and those who used multiple focused drugs at their reference date. To further align with the biosimilar introduction analysis for each hospital, we limited the analysis to patients whose reference date was between 4/1/2013 and 3/31/2018, after the year of biosimilar implementation, hospitals that had at least one sample patient in each year in the fiscal years 2013-2017, hospitals that had data in the Hospital Yearbook (*Byoin Nenkan*) 2014 (which we used for hospital characteristics), and hospitals that used at least one of the focused drugs. The final sample included 184,954 patients and 762 hospitals.

Table 2 Panel (a) shows summary statistics for the patients in the sample. They are, on average, relatively old, but they also include the young, the infants, and the babies. Males are slightly more prevalent (56%). Approximately 68% of the sample patients received chemotherapy. Only about 15% of patients have the ICD-10 code for neutropenia. Panel (b) compares the summary statistics of the patients who use at least one of the focused drugs ("Any Drug") and those who do not use any focused drugs ("Outside"). "Any Drug" patients are slightly younger, and 93.9% of them undergo chemotherapy.

Table 3 shows the number of patients for each product for each year.[2] Since the launch of biosimilars, the market share of original drugs of Filgrastim has been gradually decreasing and that of biosimilars has been increasing. Moreover, among biosimilars, there is a considerable difference in market share among companies.

Table 4 shows summary statistics for the hospitals in the sample. The biosimilar introduction rate varies slightly with their area but largely with their owner. The number of

[2]Patients who used more than one product in the data (34.3% of patients who used any of the focused drugs) were counted as the first product used.

Table 2: Summary statistics of the sample patients

(a) All

| | Mean | Min | Max | Median | Std. Dev. |
|---|---|---|---|---|---|
| Age | 66.21 | 0 | 105 | 70 | 17.20 |
| Sex (0:male 1:female) | 0.44 | 0 | 1 | - | - |
| Body Surface Area ($m^2$) | 1.56 | 0.14 | 3.18 | 1.56 | 0.23 |
| Chemotherapy | 0.68 | 0 | 1 | - | - |
| D70 | 0.15 | 0 | 1 | - | - |

(b) Any drug or Outside

| | Any Drug | | Outside | |
|---|---|---|---|---|
| | Mean | Median | Mean | Median |
| Age | 64.65 | 68 | 67.47 | 71 |
| Sex (0:male 1:female) | 0.446 | - | 0.440 | - |
| Body Surface Area ($m^2$) | 1.562 | 1.565 | 1.556 | 1.561 |
| Chemotherapy | 0.939 | - | 0.469 | - |
| D70 | 0.290 | - | 0.043 | - |
| $N$ | 82,436 | | 102,518 | |

Notes: Panel (a) shows the summary statistics of all 184,954 sample patients and Panel (b) shows that of the patients who used at least one of the focused drugs ("Any Drug") and those who do not use any focused drugs ("Outside"). Body surface area was calculated from the patients' weights and heights using the Du Bois Method (body surface area ($m^2$) = $\text{height(cm)}^{0.725} \times \text{weight(kg)}^{0.425} \times 0.007184$). "D70" is an indicator that takes a value of 1 if at least one of the disease codes (occasion of admission, comorbidity on admission, post-hospitalization onset disease, etc.) included D70, and 0 otherwise.

Table 3: Number of Patients: Drug and Fiscal Year

| Drug \FY | 2013 | 2014 | 2015 | 2016 | 2017 | Total |
|---|---|---|---|---|---|---|
| Filgrastim (original) S75$\mu$g | 6,697 | 5,819 | 3,691 | 2,449 | 1,508 | 20,164 |
| Filgrastim (original) S150$\mu$g | 851 | 710 | 475 | 325 | 226 | 2,587 |
| Filgrastim (original) S300$\mu$g | 816 | 884 | 658 | 591 | 561 | 3,510 |
| Filgrastim (original) I75$\mu$g | 208 | 214 | 113 | 74 | 40 | 649 |
| Filgrastim (original) I150$\mu$g | 51 | 34 | 24 | 37 | 23 | 169 |
| Filgrastim (original) I300$\mu$g | 123 | 111 | 82 | 73 | 48 | 437 |
| Filgrastim (F) 75$\mu$g | 46 | 546 | 651 | 929 | 1,021 | 3,193 |
| Filgrastim (F) 150$\mu$g | 1 | 46 | 70 | 108 | 87 | 312 |
| Filgrastim (F) 300$\mu$g | 13 | 54 | 86 | 72 | 86 | 311 |
| Filgrastim (Mochida) 75$\mu$g | 351 | 1,510 | 2,616 | 4,086 | 4,279 | 12,842 |
| Filgrastim (Mochida) 150$\mu$g | 22 | 132 | 226 | 376 | 320 | 1,076 |
| Filgrastim (Mochida) 300$\mu$g | 27 | 79 | 111 | 144 | 176 | 537 |
| Filgrastim (NK) 75$\mu$g | 115 | 617 | 1,131 | 1,325 | 1,529 | 4,717 |
| Filgrastim (NK) 150$\mu$g | 8 | 57 | 116 | 160 | 158 | 499 |
| Filgrastim (NK) 300$\mu$g | 17 | 43 | 51 | 40 | 56 | 207 |
| Filgrastim (Sandoz) 75$\mu$g | 0 | 0 | 121 | 220 | 218 | 559 |
| Filgrastim (Sandoz) 150$\mu$g | 0 | 1 | 12 | 12 | 5 | 30 |
| Filgrastim (Sandoz) 300$\mu$g | 0 | 0 | 15 | 21 | 4 | 40 |
| Filgrastim (Teva) 75$\mu$g | 16 | 20 | 22 | 71 | 72 | 201 |
| Filgrastim (Teva) 150$\mu$g | 1 | 0 | 1 | 7 | 2 | 11 |
| Filgrastim (Teva) 300$\mu$g | 0 | 0 | 0 | 0 | 0 | 0 |
| Lenograstim 50$\mu$g | 293 | 224 | 188 | 211 | 149 | 1,065 |
| Lenograstim 100$\mu$g | 4,556 | 4,685 | 3,769 | 3,383 | 2,393 | 18,786 |
| Lenograstim 200$\mu$g | 941 | 947 | 873 | 820 | 732 | 4,313 |
| Mirimostim 400 | 0 | - | - | - | - | - |
| Mirimostim 800 | 115 | 100 | 80 | 89 | 39 | 423 |
| Nartograstim 25$\mu$g | 4 | 8 | 7 | 8 | 4 | 31 |
| Nartograstim 50$\mu$g | 357 | 283 | 250 | 218 | 187 | 1,295 |
| Nartograstim 100$\mu$g | 21 | 10 | 9 | 7 | 8 | 55 |
| Nartograstim 250$\mu$g | 3 | 0 | 4 | 1 | 0 | 8 |
| Pegfilgrastim 3600$\mu$g | 0 | 261 | 1,829 | 1,174 | 1,145 | 4,409 |
| Outside | 18,620 | 20,221 | 20,215 | 21,255 | 22,207 | 102,518 |
| Total | 34,273 | 37,616 | 37,496 | 38,286 | 37,283 | 184,954 |

Notes: Each patient is counted as their first prescription drug. "Filgrastim (original)" refers to the original filgrastim product produced by Kyowa Hakko Kirin, and other "Filgrastim (...)" refers to the biosimilar products of filgrastim, with the names of the producing companies in parentheses.

Table 4: Summary Statistics of the sample hospitals

(a) Area

| Area | *N* | Introduce Biosimilar |
|---|---|---|
| Hokkaido | 43 | 32 (74%) |
| Tohoku | 61 | 46 (75%) |
| Kanto | 165 | 117 (71%) |
| Chubu | 131 | 95 (73%) |
| Kinki | 141 | 100 (71%) |
| Chugoku | 56 | 39 (70%) |
| Shikoku | 31 | 22 (71%) |
| Kyushu | 134 | 95 (71%) |
| Total | 762 | 546 (72%) |

(b) Owner

| Owner | *N* | Introduce Biosimilar |
|---|---|---|
| National | 87 | 77 (89%) |
| Public Medical | 297 | 225 (76%) |
| Social Insurance | 17 | 10 (59%) |
| Medical Corp. | 167 | 90 (54%) |
| Other | 82 | 50 (62%) |
| National Univ. | 48 | 43 (90%) |
| Private Univ. | 64 | 51 (80%) |
| Total | 762 | 546 (72%) |

(c) Number of Beds

| Number of Beds | Mean | Min | Max | Median | Std. Dev. | *N* |
|---|---|---|---|---|---|---|
| Introduce Biosimilar | 480.94 | 80 | 1,505 | 435 | 233.17 | 546 |
| No Biosimilar | 340.97 | 60 | 1,275 | 279 | 233.82 | 216 |
| Total | 441.27 | 60 | 1,505 | 400 | 238.91 | 762 |

Notes: All panels show summary statistics of the 762 sample hospitals."Introduce Biosimilar" is defined as hospitals that prescribed biosimilars (in our focused drugs) at least once during the sample period.

beds in the biosimilar introduction hospitals is larger than other hospitals on average.

# 4 Hospital Model of Biosimilar Introduction

For pharmaceutical products used in hospitals, such as the focused drugs examined in this study, each hospital decides in advance which pharmaceuticals to use and manage. Figure 2 Panel (a) shows the share of the focused drug combinations used in FY2017 by hospitals in our dataset. There are clear differences in the types of drugs used by different hospitals. Panel (b) focuses on Filgrastim and shows that while 44.09% of hospitals prescribed only biosimilars in FY2017, 14.17% of them prescribed only brand name drugs, and 25.46% of them prescribed both, indicating that hospitals use different strategies.

Although the launch date of biosimilars is the same throughout Japan, the timing of their introduction varies significantly between hospitals. Figure 3 shows the proportion of hospitals that introduce biosimilars monthly. Although the adoption rate defined by the first biosimilar prescription (red line with circles) at the end of the sample period (March 2018) was 72%, the timing of adoption differed for each hospital and biosimilars were gradually spreading. If we define adoption by the first biosimilar prescription and final prescription of the original drug in the sample period (blue line with triangles), this tendency remains.[3]

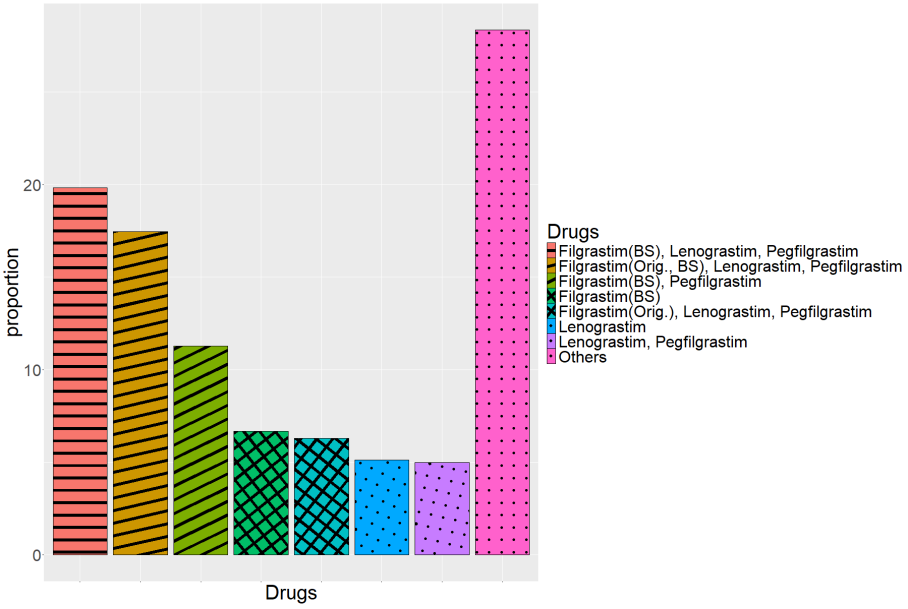
Therefore, before considering a model for patients' choice of drugs, we first consider the model of biosimilar adoption by hospitals. We construct a simple theoretical model that describes biosimilar adoption in Section 4.1 and then the empirical model in Section 4.2. Section 4.3 presents the estimation results.

## 4.1 Theoretical Model

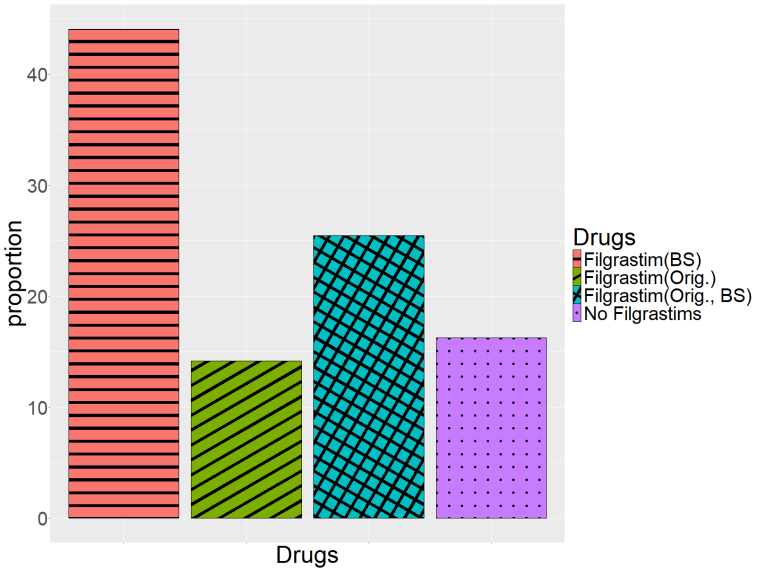
When a biosimilar enters the market, each hospital decides whether to use it, considering factors such as its efficacy, drug price, wholesale price, and administrative costs. Here we construct a simple theoretical hospital model that describes the biosimilar adoption in our sample. Hospitals choose whether to prepare a original or biosimilar drug based on their profits and patient benefits. For simplicity, we assume that the number of patients in each hospital is fixed and normalized to one. For each patient, she and/or her doctor choose whether to use the pharmaceutical prepared by the hospital (original or biosimilar) or nothing (cost 0). For the institutional reasons explained in Section 3, the sale for each patient is

[3]Of the 546 hospitals that prescribed biosimilars at least once, 397 (72.71%) hospitals also prescribed biosimilars in the last month of the data period (March 2018), and 487 (89.19%) hospitals had their last prescription in the data period after December 2017, suggesting that the majority continued to adopt biosimilars during the data period.

Figure 2: Share of drug combinations used in FY2017 by hospital



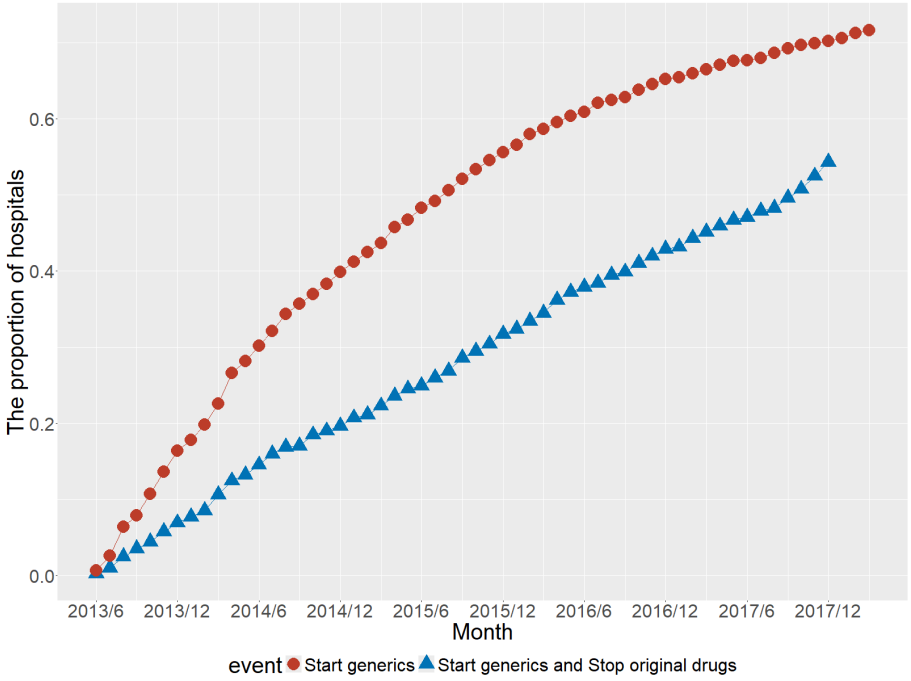

(a) All Drugs




(b) Filgrastim

Notes: These graphs show the share of hospitals categorized by drugs prescribed in FY2017. The use of each drug in each hospital was judged based on whether the drug was prescribed at least once in that hospital in FY2017.

Figure 3: Proportion of hospitals introducing biosimilars in each month




Notes: The red line (circles) shows the proportion of hospitals introducing biosimilars where the introduction of biosimilars was defined as the first biosimilar prescription in that hospital. The blue line (triangles) shows the proportion where the introduction was defined by the first biosimilar prescription and final prescription of the original drug (in the sample period) in that hospital. We omitted the blue line in 2018 because it rapidly became similar to the red line by definition.

fixed (total sales are fixed), and only the wholesale prices of the drugs are considered as their costs. Then the revenue function of the hospital that chooses drug $j$ ($\in \{o, b\}$, corresponding to "original" and "biosimilar," respectively) is

$$\pi_j = R - s_j(p_j)p_j + EU_j(p_j) - C_j \quad (j = o, b) \tag{2}$$

where $R$ is the total sales (fixed), $p_j$ is the wholesale price of drug $j$, $s_j(p_j)$ is the share of patients who use drug $j$ (and $1 - s_j(p_j)$ patients use nothing), $EU_j(p_j)$ is ex-ante average utility of patients if the hospital chose drug $j$ the hospital takes into account, and $C_j$ is the administrative costs when the hospital prepares drug $j$.

The hospital chooses drug $b$ if the revenue from that is larger than that from a counterpart, that is, if $\pi_b \geq \pi_o \Leftrightarrow \pi_b - \pi_o \geq 0$. To consider the relationship between prices and biosimilar introduction, we differentiate $\pi_b - \pi_o$ with respect to $p_o$ and $p_b$ and have

$$\frac{\partial(\pi_b - \pi_o)}{\partial p_o} = -\frac{\partial EU_o(p_o)}{\partial p_o} + s_o(p_o) + \frac{\partial s_o(p_o)}{\partial p_o}p_o \tag{3}$$

$$\frac{\partial(\pi_b - \pi_o)}{\partial p_b} = \frac{\partial EU_b(p_b)}{\partial p_b} - s_b(p_b) - \frac{\partial s_b(p_b)}{\partial p_b}p_b. \tag{4}$$

The first term in Equation (3) indicates the decrease in the patients' average utility, the second term indicates the increased cost due to the price change, and the third term represents the decreasing cost due to the decreasing share of patients who use drug $a$. If the third term effect is dominated by the first and second term effects, $\partial(\pi_b - \pi_o)/\partial p_o > 0$ holds and then the increase of $p_o$ increases the biosimilar introduction and vice versa. The interpretation of Equation (4) is the same and the only difference is their sign. If the third term effect is dominated by the first and second term effects, $\partial(\pi_b - \pi_o)/\partial p_b < 0$ holds and then the increase of $p_b$ decreases the biosimilar introduction and vice versa.[4]

In the simple setting, we consider each hospital decides which drug will they use in each period given the wholesale prices. Adding hospital-specific switching costs to this simple model and considering the hospital-specific time needed to switch can explain the gradual spread of biosimilars (Figure 3). Other explanations include the gradual reduction of uncertainty about the efficacy and risk of biosimilars. However, the important point is that, in all cases, wholesale prices are factors that can affect hospital decision-making, and the decisions are made each period. To incorporate these features, we estimate the hazard

[4]We can easily extend this model to allow hospitals to prepare for both original drugs and biosimilars. In that case, Equation (4) remains very similar and almost the same proposition holds. Contrastingly, Equation (3) becomes slightly more complicated because the change in $p_o$ also affects the share of each drug and patients' average utility in the revenue function, and these are not erased when we differentiate the revenue by $p_o$.

model and include wholesale prices as explanatory variables to analyze hospitals' biosimilar introduction behavior.

## 4.2 Estimation Model

We estimated the effect of the drug prices of original products and biosimilars on the timing of biosimilar introduction in each hospital using the time-varying Weibull hazard model. We represent the probability that the number of days until the introduction of a biosimilar is greater than $t$ as a survival function $S(t) = \Pr(T > t)$. We define

$$S(t_h) = \exp(-\lambda(t_h) \cdot t_h^p) \tag{5}$$

$$\lambda(t_h) = \exp(\alpha_1 p_{\text{original},t_h} + \alpha_2 p_{\text{biosim},t_h} + \beta x_h) \tag{6}$$

where $h$ indicates hospital, $t_h$ indicates the number of days from biosimilar launch (May 31, 2013) to biosimilar adoption by hospital $h$, $p_{\text{original},t_h}$ indicates the wholesale price of the original drug, $p_{\text{biosim},t_h}$ indicates that of biosimilars, $x_h$ is a characteristic vector of hospital $h$, and $\alpha, \beta,$ and $p$ are parameters which will be estimated ($p$ is a "shape parameter" of this hazard model). The estimated values are parameters that maximize the likelihood calculated from $S(t_h)$. This model can be transformed into an easily interpretable form as follows:

$$\log(t_h) = \alpha_1 p_{\text{original},t} + \alpha_2 p_{\text{biosim},t} + \beta x_h + \epsilon_h \tag{7}$$

where $\epsilon_h$ is an independent and identically distributed shock following an extreme distribution. This equation indicates that positive (negative) coefficients of price indicate that the higher price delays (accelerates) the biosimilar introduction, and the expected time to introduction becomes longer (shorter) on average. Specifically, we use the wholesale price of Filgrastim (original) Syringe 75$\mu$g (the most prevalent original drug product) as a $p_{\text{original},t_h}$ and that of Filgrastim (Mochida) 75$\mu$g (the most prevalent biosimilar product) as a $p_{\text{biosim},t}$. The characteristics of the hospital we used are $x_h = (\boldsymbol{area}_h, \boldsymbol{owner}_h, bed_h)$ where $\boldsymbol{area}_h$ is a region dummy (which divides Japan into eight regions), $\boldsymbol{owner}_h$ is a owner dummy (seven categories basically defined by the Ministry of Health, Labour, and Welfare, Japan), and $bed_h$ is the number of beds. The sample included 762 hospitals for which sample patients were present every year during the sample period. To determine the average effect per patient, each hospital was weighted by the total number of sample patients in the sample period.

One concern is that the wholesale prices ($p_{\text{original},t_h}$, $p_{\text{biosim},t_h}$) used in this analysis are the national average calculated from the Equation (1) but the actual wholesale prices may

Table 5: Estimation Results: Weibull Hazard Model of Biosimilar Introduction

| | (1) All | (2) Larger Hospitals |
|---|---|---|
| $p_{\text{original},t}$ | -0.0042644* | -0.0043306* |
| | (0.0021912) | (0.0022993) |
| $p_{\text{biosim},t}$ | 0.0035618* | 0.0036505* |
| | (0.0019861) | (0.0020872) |
| # of Beds | -0.0001282 | 0.0003747 |
| | (0.0002488) | (0.0002818) |
| ln(p) (Shape Parameter) | 0.1894726** | 0.2038324** |
| | (0.0847713) | (0.0919238) |
| Region Dummies | ✓ | ✓ |
| Owner Dummies | ✓ | ✓ |
| $N$ | 762 | 381 |
| log-pseudo likelihood | -271565.26 | -245139.97 |

Notes: The unit of observation is the hospital. $p_{\text{original},t}$ and $p_{\text{biosim},t}$ are the wholesale prices of Filgrastim (original) Syringe 75$\mu$g and Filgrastim (Mochida) 75$\mu$g calculated from Equation (1). Each observation was weighted by the total number of sample patients at that hospital in the sample period. The robust standard errors are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively.

differ from hospital to hospital. For example, a large hospital may purchase a larger volume of pharmaceuticals and thus receive a discount. However, because the model controls for hospital characteristics ($x_h = (\boldsymbol{area}_h, \boldsymbol{owner}_h, bed_h)$), this does not cause a serious estimation problem as long as the actual wholesale prices for each hospital can be sufficiently explained by these variables.

## 4.3 Results

Column (1) of Table 5 presents the Weibull hazard regression results. These results are consistent with the implication that the higher the wholesale price of the original product is, the faster the introduction of the biosimilar is, and the higher the wholesale price of the biosimilar is, the slower its introduction is.

One concern is that the timing of biosimilar introduction may not have been correctly measured for hospitals with few patients because the timing of introduction was defined as the first time some biosimilar products were prescribed to a patient in that hospital.

Therefore, we focused on larger hospitals with a median number of sample patients or more and conducted the same analysis. These hospitals accounted for more than 90% of all sample patients. Column (2) of Table 5 shows the results using only larger hospitals; these results are very similar to those obtained using all hospitals.

Based on these results, we can simulate the timing of biosimilar introduction for each hospital in the counterfactual scenarios when the prices of the original and biosimilar products are changed. By combining this with the patient-level drug choice model to be estimated next, we can also calculate biosimilar usage rates and market shares for the counterfactual cases.

# 5 Patient Model of Drug Choice

## 5.1 Model

In this section, we consider a pharmaceutical choice model for sample patients. As this choice is affected (or may be determined in many cases) by doctors in the hospital, this model includes not only patient' characteristics but also variables related to doctors and hospitals, such as the wholesale prices of drugs (Section 3). The majority (65.70%) of patients who used focused drugs used only one product, therefore, we considered a cross-sectional static discrete choice model, where we defined the first drug used by a patient as their choice.

Patient $i$ on the reference date $s$ in FY $t$ in hospital $h$ selects drug (or outside option) $j$ from their choice set $J_{h(i),s}$ that maximizes their utility $u_{ijt}$

$$\max_{j \in J_{h(i),s}} u_{ijt} = (\alpha_0 + \boldsymbol{\alpha}_1 \boldsymbol{z}_i) p_{jt} + \xi_j + \boldsymbol{\beta}_1 \boldsymbol{z}_i \boldsymbol{x}_j + \omega_0 \ln(month_{js} + 1) + \epsilon_{ijt} \tag{8}$$

where $p_{jt}$ indicates the wholesale price, $x_j = (drug_j, generic_j, \boldsymbol{molecule}_j)$ indicates pharmaceutical characteristics where $drug_j$ is a drug (not outside option) dummy, $generic_j$ is a generic dummy, and $\boldsymbol{molecule}_j$ is a molecule dummy vector, $\xi_j$ is an unobservable drug characteristic (scalar), $z_i = (z_{1i}, z_{2i})$ are observable patient characteristics where $z_{1i}$ includes age, age$^2$, sex, body surface area, a chemotherapy dummy, and $z_{2i}$ includes a neutropenia code (D70) dummy, $month_{js}$ indicates months after launch, $\epsilon_{ijt}$ is an independent and identically distributed shock following extreme distribution, and $(\alpha_0, \alpha_1, \beta_1, \gamma_0, \gamma_1, \omega_0) \equiv \theta_D$ are parameters which will be estimated.

The choice set $J_{h(i),s}$ is the set of drugs available at hospital $h(i)$ for patient $i$ on reference date $s$ (including the outside option) and is constructed as follows: For each pair of hospital $h$-drug $j$, we used all available data (not limited to the sample patients) to determine whether

the drug was prescribed, and the first and last prescription dates if it was prescribed at least once. If the patient $i$'s hospital $h(i)$ had prescribed drug $j$ at least once and reference date $s$ was within the range of its first and last prescription dates, then $j \in J_{h(i),s}$, otherwise, $j \notin J_{h(i),s}$. The 689 patients for whom $J_{h(i),s}$ included only an outside option were excluded from the estimation.

Due to computational difficulties, we imposed several assumptions on the parameters and estimated the model

$$\max_{j \in J_{h(i),s}} u_{ijt} = \delta_{jt} + \boldsymbol{\alpha}_1 \boldsymbol{z}_{1i} p_{jt} + \beta_{11} z_i drug_j + \beta_{12} z_{1i} generic_j + \beta_{13} z_{2i} molecule_j + \omega_0 \ln(month_{js} + 1) + \epsilon_{ijt} \tag{9}$$

where $\delta_{jt}$ is a mean utility of drug $j$ in FY $t$:

$$\delta_{jt} \equiv \alpha_0 p_{jt} + \xi_j. \tag{10}$$

## 5.2 Estimation Method

Our estimation method is based on Dunn (2012)'s method. At the first stage, we estimated the following equation using the maximum likelihood method, where $\boldsymbol{\alpha}_1$, $\boldsymbol{\beta}_1$, $\omega_0$, and $\delta_{jt}$ were identified.

$$\Pr_{it}(j|\boldsymbol{\delta}, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \frac{\exp(\delta_{jt} + \boldsymbol{\alpha}_1 \boldsymbol{z}_{1i} p_{jt} + \beta_{11} z_i drug_j + \beta_{12} z_{1i} generic_j + \beta_{13} z_{2i} molecule_j)}{\sum_{k \in J_{hs}} \exp(\delta_{kt} + \boldsymbol{\alpha}_1 \boldsymbol{z}_{1i} p_{kt} + \beta_{11} z_i drug_k + \beta_{12} z_{1i} generic_k + \beta_{13} z_{2i} molecule_k))} \tag{11}$$

In the second stage, $\hat{\delta_{jt}}$ estimated at the first stage was regressed on the wholesale price and each product dummy to obtain the remaining parameter $\alpha_0$.

$$\hat{\delta_{jt}} = \alpha_0 p_{jt} + \sum_{j'=1}^{J} \xi_{j'} \mathbf{1}\{j = j'\} + e_{jt} \tag{12}$$

In this regression, the wholesale price $p_{jt}$ may have been correlated with the error term $e_{jt}$ and the ordinary least squares estimates may be biased. This endogeneity problem could arise if, for example, wholesale prices are determined through negotiations between hospitals and pharmaceutical companies and the unobserved time-varying quality of the product affects their bargaining power. To address this problem, we used Dunn (2012)'s instrumental variable approach based on Gaynor and Vogt (2003).

The instruments we used are demand and markup predicted from the parameters es-

timated in the first stage but calculated with $\boldsymbol{\alpha}$ and $e$ set to zero. First, we defined the demand as

$$D_{jt} \equiv \sum_{i=1}^{I} \Pr_{it}(j|\boldsymbol{\delta}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \tag{13}$$

and responsiveness to prices as

$$\frac{\partial D_{jt}}{\partial p_{jt}} \equiv \sum_{i=1}^{I} \frac{\partial \Pr_{it}(j|\boldsymbol{\delta}, \boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial p_{jt}}. \tag{14}$$

Based on these definitions, we constructed the demand instrument

$$D_{jt}^{I}(j|\boldsymbol{\delta}(\boldsymbol{\alpha} = \mathbf{0}, \boldsymbol{e} = \mathbf{0}), \boldsymbol{\alpha} = \mathbf{0}, \boldsymbol{\beta}) \tag{15}$$

and markup instrument

$$\frac{D_{jt}^{I}(j|\boldsymbol{\delta}(\boldsymbol{\alpha} = \mathbf{0}, \boldsymbol{e} = \mathbf{0}), \boldsymbol{\alpha} = \mathbf{0}, \boldsymbol{\beta})}{\frac{\partial D_{jt}^{I}(j|\boldsymbol{\delta}(\boldsymbol{\alpha}=\mathbf{0},\boldsymbol{e}=\mathbf{0}),\boldsymbol{\alpha}=\mathbf{0},\boldsymbol{\beta})}{\partial p_{jt}}}. \tag{16}$$

As the generic market may have different features compared to the original drug market, we constructed additional instruments by interacting the two instruments above with a generic dummy. Finally, we used four instrumental variables.

Why are these instrumental variables valid? In the standard firm profit-maximizing model, the demand and markup of a product affect its pricing and satisfy the condition for instrumental variables correlated with price. However, demand and markups do not satisfy the exclusion restriction because they depend on the error term $e_{jt}$ directly and through price $p_{jt}$. Therefore, by substituting $\alpha = e = 0$ and removing the part of the demand and markup that depends on the price $p_{jt}$ or the error term $e_{jt}$, they become valid instrumental variables that satisfy the exclusion restriction. Intuitively, these instruments exploit the structure that the individual demographics $z_i$ of the patient are correlated with price $p_{jt}$ through demand and markup but are not correlated with the error term $e_{jt}$. Computationally, $\delta(\alpha = 0, e = 0)$ is calculated as the average of $\hat{\delta_{jt}}$ estimated in the first stage for product $j$. This is because if we impose $\alpha = 0$ and $e = 0$ in the second stage Equation (12); only product dummies that are independent of year $t$ remain.

Another endogenous concern is the bias that comes from patients' choice of hospital. For example, if patients who prefer original drugs choose to be admitted to hospitals that prescribe original drugs, the $\delta_{jt}$ of original drugs are overestimated because the share of

original drugs in these hospitals is higher than average. However, since the drugs under study are usually used to reduce the side effects of cancer treatment, it is unlikely that patients choose hospitals with these drugs in mind; thus, the bias is considered to be small.

### 5.3 Results

Table 6 presents the results of the first stage of the maximum likelihood estimation. Many coefficients have statistically significant estimates, some of which are reasonable. For example, the positive coefficient for Drug×Pchemo indicates that patients using chemotherapy are more likely to use one of the drugs, and a positive coefficient for Drug×D70 indicates that patients diagnosed with neutropenia are also more likely to use one of the drugs.

Table 7 presents the first-step regression results of the second stage, which regresses the wholesale price on the instruments. The demand instrument and its interaction term with the generic dummy variable were strictly significant. Table 8 presents the results of the second-stage regression. In all the cases, the price effect was not significant. This may reflect the fact that given hospital-prepared drugs by the choice set $J_{h(i),s}$, wholesale prices are no longer considered when the patient or doctor makes a choice.

## 6 Simulation

To what extent does a drug price change in original products and biosimilars promote the diffusion of biosimilars? This is a key question to consider in pharmaceutical price-setting policies aimed at promoting the use of biosimilars and reducing medical costs. In this section, we conduct simulations to answer this question quantitatively by combining the hospital model of biosimilar introduction in Section 4 and the patient model of drug choice in Section 5 ("the hospital-set model"). In addition, we estimate and simulate a model that does not use the hospital model but only the patient model and does not have a product choice set for each hospital (common to all hospitals)("the all-set model"), and compare the results with those of the hospital-set model. First, we compared the simulated biosimilar shares at the actual price to the actual shares and then calculated the simulated share when the price was changed hypothetically as a counterfactual scenario.

### 6.1 Simulation Procedure

The simulation procedures for the hospital-set model are as follows: First, based on the hazard model estimation results described in Section 4, we calculated the predicted survival curve for biosimilar introduction for each hospital at the product prices specified in the

Table 6: First Stage Multinomial Logistic Regression Results

| | |
|---|---|
| Price×Age | 0.000000168** |
| | (6.70e-08) |
| Price×Age$^2$ | -4.29e-10 |
| | (5.93e-10) |
| Price×Sex | 0.00000140*** |
| | (0.000000480) |
| Price×Bsa | 0.00000962*** |
| | (0.00000126) |
| Price×Pchemo | 0.00000666*** |
| | (0.000000947) |
| Drug×Age | 0.0430*** |
| | (0.00219) |
| Drug×Age$^2$ | -0.000487*** |
| | (0.0000203) |
| Drug×Sex | -0.0752*** |
| | (0.0176) |
| Drug×Bsa | -0.805*** |
| | (0.0452) |
| Drug×Pchemo | 2.540*** |
| | (0.0236) |
| Drug×D70 | 1.827*** |
| | (0.0385) |
| Filgrastim×D70 | 0.209*** |
| | (0.0354) |
| Nartograstim×D70 | 0.987*** |
| | (0.0661) |
| Lenograstim×D70 | 0.275*** |
| | (0.0370) |
| Generic×Age | 0.0191*** |
| | (0.00306) |
| Generic×Age$^2$ | -0.0000517* |
| | (0.0000275) |
| Generic×Sex | 0.0211 |
| | (0.0230) |
| Generic×Bsa | 0.224*** |
| | (0.0608) |
| Generic×Pchemo | 0.213*** |
| | (0.0352) |
| ln(month + 1) | 0.0577 |
| | (0.0499) |
| product×year dummies | ✓ |
| $N$ | 184,265 |

Notes: "Price" indicates the wholesale price calculated using Equation (1). Body surface area was calculated from the patients' weight and height using the Du Bois Method (body surface area ($m^2$) = height(cm)$^{0.725}$ × weight(kg)$^{0.425}$ × 0.007184). "D70" is an indicator that takes the value of 1 if at least one of the disease codes (occasion of admission, comorbidity on admission, or post hospitalization onset disease, etc.) included D70, and 0 otherwise. The robust standard errors are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively.

Table 7: Second Stage Instrumental Variable Estimation Results: first step

| | (2) | (3) |
|---|---|---|
| | price | price |
| Demand IV | 0.683*** | 0.498** |
| | (0.178) | (0.250) |
| Demand IV× Generic | -1.514*** | -1.118*** |
| | (0.268) | (0.295) |
| Markup IV | - | -0.463 |
| | | (0.449) |
| Markup IV × Generic | - | 0.813* |
| | | (0.460) |
| product dummies | ✓ | ✓ |
| $N$ | 134 | 134 |
| adj. R-squared | 0.993 | 0.993 |
| F test | F(2, 103) = 15.91 | F(4, 101) = 18.18 |
| Prob > F | 0.0000 | 0.0000 |

Notes: We use only those product $j$–year $t$ pairs for which $\delta_{jt}$ was precise estimated where the z-values are greater than 2 in the first stage in Table 6. "Price" indicates the wholesale price calculated from Equation (1). "Demand IV" is defined by Equation (15) and "Markup IV" is defined by Equation (16). The standard errors are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively.

Table 8: Second Stage Instrumental Variable Estimation Results

| | (1) | (2) | (3) |
|---|---|---|---|
| | | product×year dummies | |
| | OLS | IV(Demand) | IV(Demand, Markup) |
| price | 0.0000269 | 0.00000456 | -0.0000228 |
| | (0.0000364) | (0.0000581) | (0.0000694) |
| product dummies | ✓ | ✓ | ✓ |
| $N$ | 134 | 134 | 134 |
| adj. R-squared | 0.888 | 0.887 | 0.885 |

Notes: We used only those product $j$–year $t$ pairs for which $\delta_{jt}$ was precisely estimated where the z-values were greater than 2 in the first stage in Table 6. "Price" indicates the wholesale price calculated using Equation (1). Column (1) shows the results of ordinary least squares regression. Column (2) shows the results of the instrumental variable regression, which uses "Demand IV" defined by Equation (15) and the interaction of "Demand IV" and a generic dummy as instruments. Column (3) shows the results of the instrumental variable regression, which also uses "Markup IV" defined by Equation (16) and the interaction of "Markup IV" and a generic dummy as instruments, in addition to the instruments in column (2). The standard errors are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively.

scenario. Second, for each patient, we obtained the predicted probability of biosimilar introduction to their hospital on their reference date. Third, using this probability as a weight, each patient was split into pre- and post-biosimilar-introduction observations. For example, if the predicted biosimilar-introduction probability was 0.3, the weight for a pre-biosimilar observation was 0.7 and that for a post-biosimilar observation was 0.3. Fourth, a counterfactual pharmaceutical choice set was constructed for each observation. For pre-biosimilar observations, we excluded all biosimilars from the (original) choice set of the hospital on the reference date. For post-biosimilar observations, we added to the (original) choice set on the reference date all biosimilars that the hospital had prescribed at least once in the original data and that had already been launched at the reference date. For hospitals that had not prescribed any biosimilars in the original data, we added the biosimilar product with the largest market share, Filgrastim (Mochida) 75$\mu$g, to the choice set. Fifth, based on the patient model in Section 5, we calculated the choice probability for each product for each observation using the product price specified in the scenario, the choice set modified in the fourth step, and the estimated parameters of the patient model. Finally, the calculated choice probabilities were weighted and summed to obtain the number of patients using each product in each year.

To simulate the all-set model, we calculated the choice probability of each product for each patient based on the product price specified in the scenario and the estimated parameters of the patient model with the choice set that included all products available at the reference date and was common to all hospitals. The choice probability was then summed to obtain the number of patients using each product in each year.

## 6.2 Actual Scenario

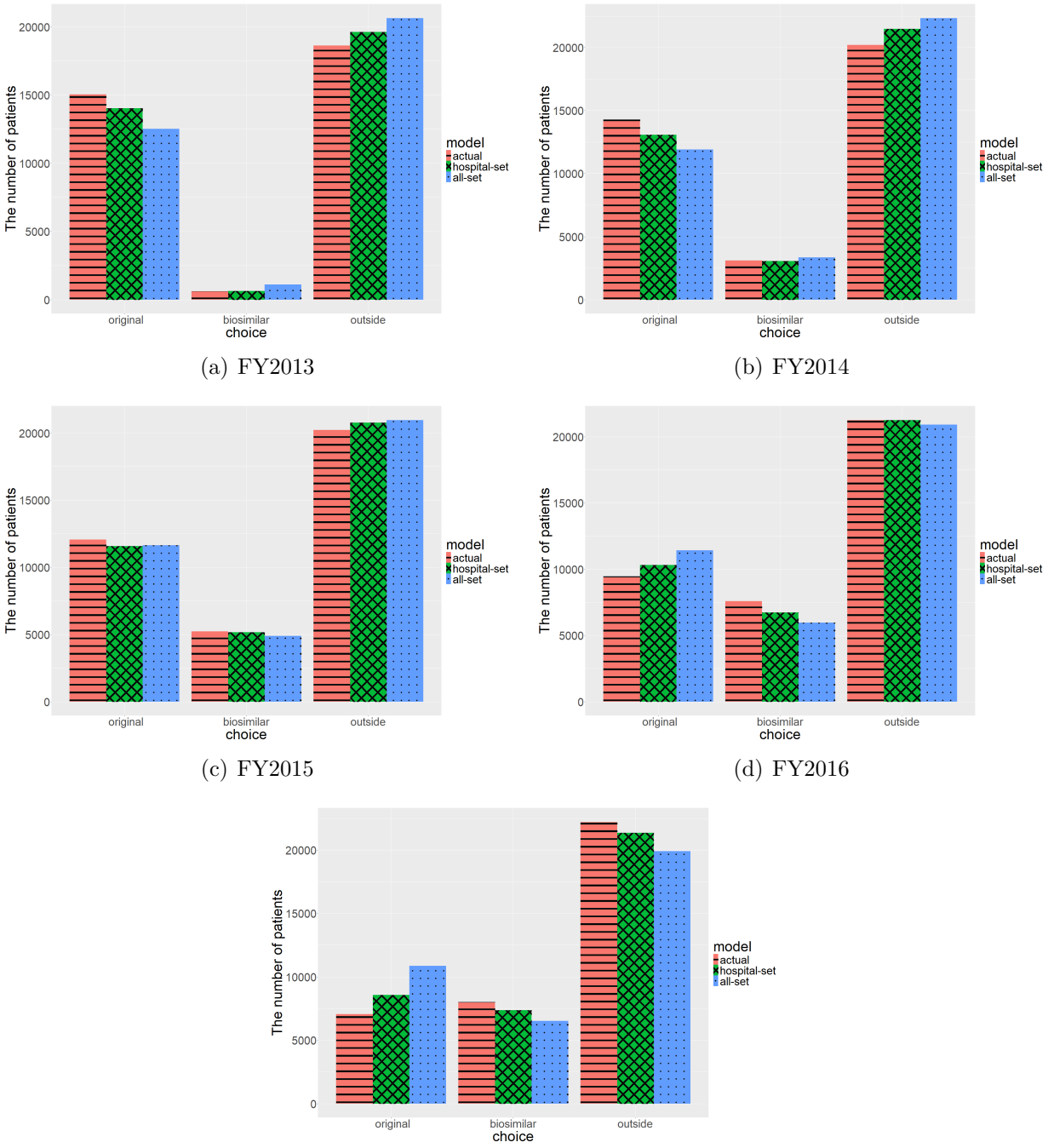
Here, we simulate the actual scenario, that is, the product prices are set to the actual prices, and the simulated results are compared with the actual number of patients for each year.

Figure 4 shows the simulation results of the hospital-set model, simulation results of the all-set model, and actual share of the sample data. In almost all years, the hospital-set model predicted values closer to the actual share than the all-set model. Particularly, in FY2016 and FY2017, the hospital-set model was closer to the actual share than the single model, which overpredicted the original drug share and underpredicted the biosimilar share.

## 6.3 Hypothetical Scenario

We simulated a counterfactual scenario if the wholesale price of biosimilars was 10% lower (the prices of other drugs would remain the same) to consider the extent to which lowering

Figure 4: Simulated Share at Actual Prices




(a) FY2013

(b) FY2014

(c) FY2015

(d) FY2016

(e) FY2017

Note: These graphs show the actual and simulated numbers of patients who used original drugs and biosimilars and those who did not use any drugs in each fiscal year. The bars "actual" are the actual number of patients. The bars "hospital-set" indicate the simulation results using the estimation results of Table 5 (1), 6, and 8 (3) with actual prices. The bars "all-set" indicate the simulation results using the estimation results of Table A1 and A2 (3) with actual prices. The detailed simulation procedure is described in Section 6.1.

the price of biosimilars could promote their widespread use.

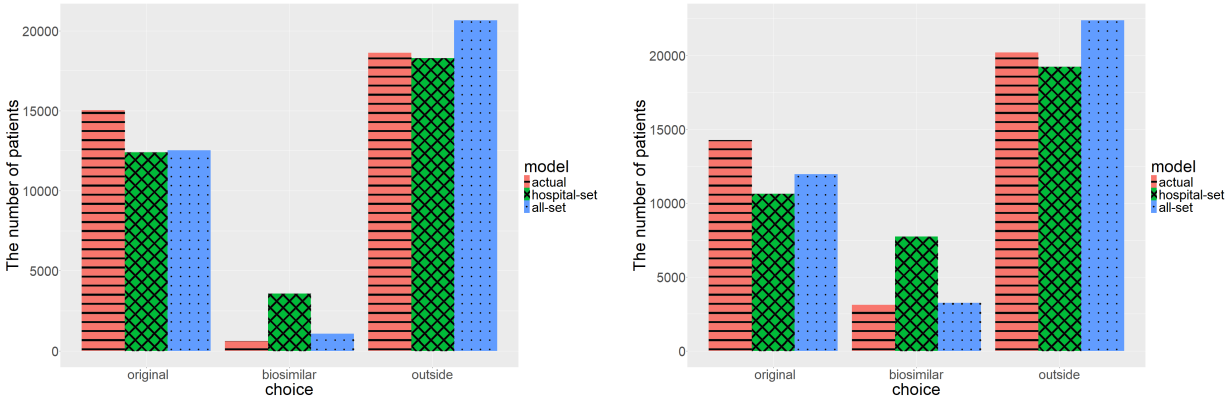
Figure 5 shows the simulation results of the hospital-set model, simulation results of the all-set model, and actual shares of the sample data. The hospital-set model showed a stronger penetration effect by lowering the price of biosimilars compared to the all-set model, with the number of patients using biosimilars increasing by 12.04% in FY2017 compared to the actual value. Particularly, in FY2016 and FY2017, the hospital-set model yielded more reasonable results than the all-set model, which yielded the non-intuitive result that the share of biosimilars was smaller than the actual value when wholesale prices decreased. Therefore, a hospital model for biosimilar introduction is important for considering and simulating pharmaceutical price changes in this market.

To confirm whether hospitals or patients drive these results, we ran the following four hypothetical scenarios and compared the results: (1) actual prices, (2) wholesale prices of biosimilars (BS) in the discrete choice model were 10% lower (the biosimilar prices in the hospital model would remain the same), and (3) wholesale prices of biosimilars in the hospital model were 10% lower (the biosimilar prices in the discrete choice model would remain the same); and (4) wholesale prices of biosimilars in both models were 10% lower. Figure 6 presents the simulation results for the four scenarios. Panel (e) shows that the share of generic users increased slightly when only the price in the discrete choice model was changed ("BS 10% lower for patients"), but when only the price in the hospital model was changed ("BS 10% lower for hospitals"), the increase in generic users and decrease in original users were quite similar to that when the biosimilar price in both models was changed ("BS 10% lower"). The other panels exhibit similar results. Therefore, the main factor responsible for the increase in the number of generics users when the generic price decreases is the hospital's decision to introduce generics. This suggests that if the government wants to promote the use of generics and reduce medical costs, a policy targeting hospitals works better than one targeting patients or doctors.

# 7 Conclusions

In the analysis of the pharmaceutical market, it is generally assumed that a patient or physician chooses a drug from all drugs available in the market at that time. However, in hospitals, the board of directors, for example, decides in advance which drugs to purchase, and the patient is prescribed those drugs. In this study, we proposed a model for drug demand that explicitly incorporates the introduction of generic drugs by hospitals. Using this model, we analyzed the market for the biologic drug Filgrastim and its generic in Japan. The simulation results showed that if the price of a generic drug were to decrease by 10%,

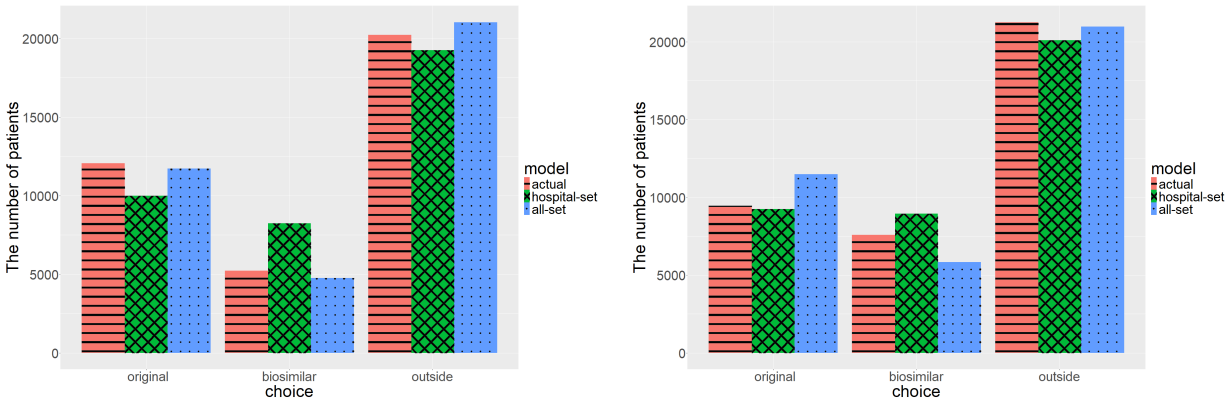
Figure 5: Simulated Share at Hypothetical Prices, 10% lower biosimilars prices




(a) FY2013



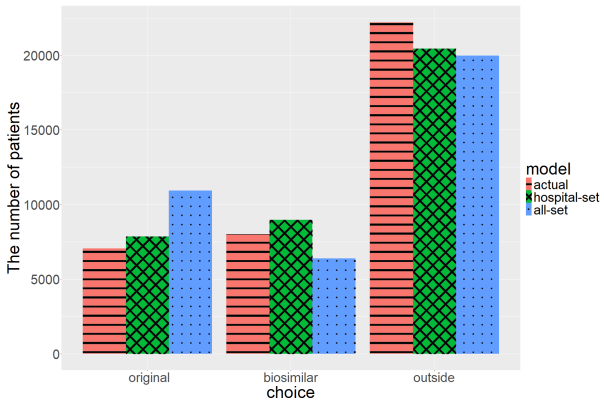

(b) FY2014




(c) FY2015



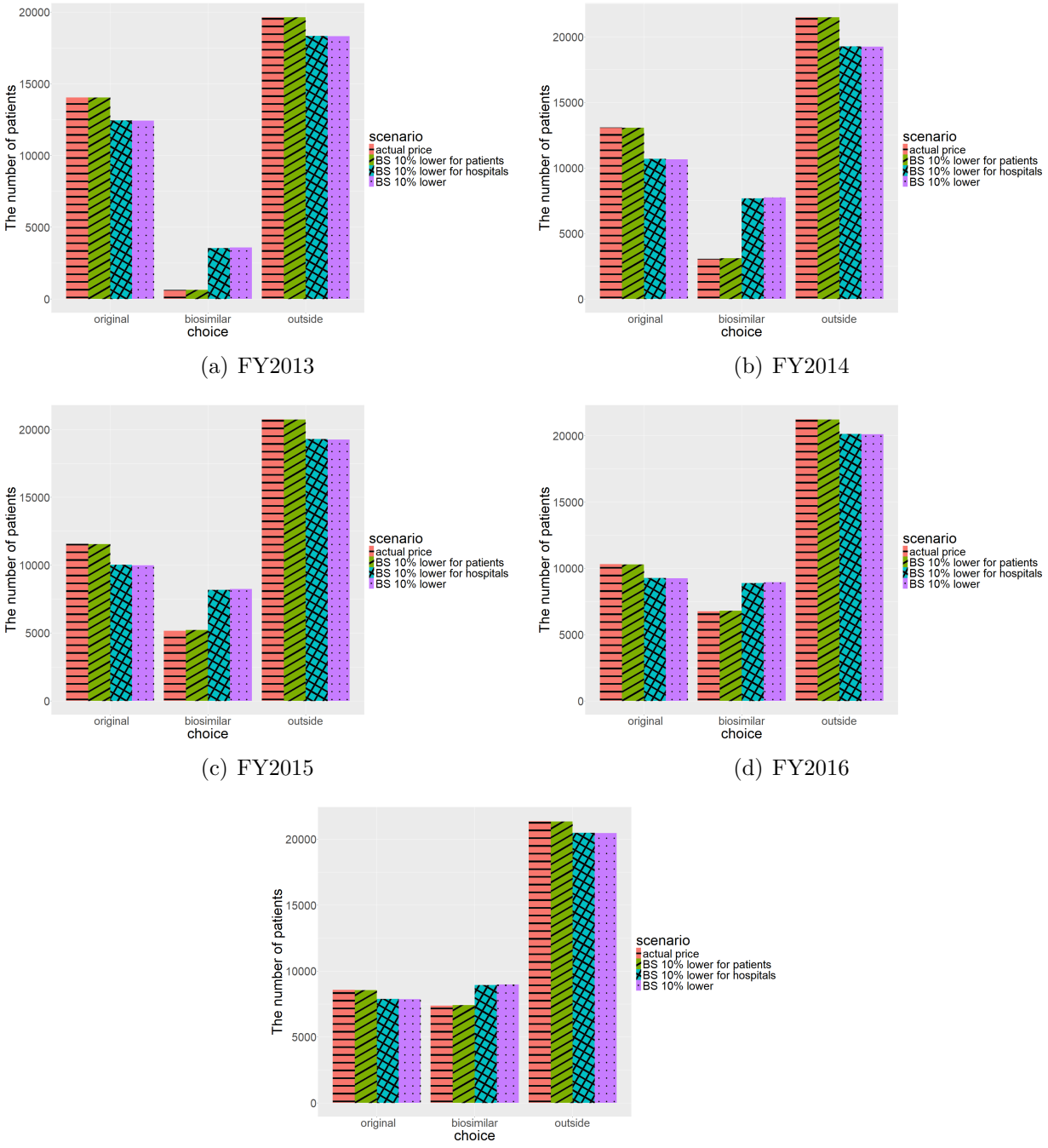

(d) FY2016




(e) FY2017

Note: These graphs show the actual and simulated numbers of patients who used original drugs and biosimilars and those who did not use any drugs in each fiscal year. The bars "actual" are the actual number of patients. The bars "hospital-set" indicate the simulation results using the estimation results of Table 5 (1), 6, and 8 (3) with 10% lower biosimilar prices. The bars "all-set" indicate the simulation results using the estimation results of Table A1 and A2 (3) with 10% lower biosimilar prices. The detailed simulation procedure is described in Section 6.1.

Figure 6: Simulated Share at Hypothetical Prices, Factor Decomposing




(a) FY2013

(b) FY2014

(c) FY2015

(d) FY2016

(e) FY2017

Note: These graphs show the simulated numbers of patients who used original drugs and biosimilars and those who did not use any drugs in each fiscal year. All simulations used the estimation results of Table 5 (1), 6, and 8 (3). The bars "actual price" are the simulation results with actual prices. The bars "BS 10% lower for patients" are the simulation results with 10% lower biosimilar prices in the discrete choice model. The bars "BS 10% lower for hospitals" show the simulation results with 10% lower biosimilar prices in the discrete choice model. The bars "BS 10% lower" represent the simulation results with 10% lower biosimilar prices for both models. The detailed simulation procedure is explained in Section 6.1.

its market share would increase by 12.04% in the most recent year, and that this spreading effect mostly comes from the hospitals' decision to introduce generics.

Our study has several limitations. First, this study analyzed a specific subject, pharmaceuticals for neutropenic patients in Japan. Therefore, it is necessary to confirm from other data for different systems and different drugs whether the model presented in this study is more reasonable than existing models.

Second, we only analyzed a limited part of the set of drugs prepared by the hospitals, namely whether biosimilars were introduced, owing to the limited price variation within the sample period. If we could estimate the entire set of drugs owned by hospitals using data for a longer period or from regions with larger price fluctuations, we could capture the impact of the policy in more detail.

Finally, although this study focused only on the analysis of drug demand, an analysis of the profit and development incentives of pharmaceutical companies is also essential to consider the impact of the drug price regulation policy. A comprehensive analysis of the profit and development incentives of pharmaceutical companies, including the data and methods to be used., is important for the future,

# References

**Berry, Steven**, "Estimating discrete-choice models of product differentiation," *The RAND Journal of Economics*, 1994, *25* (2), 242–262.

**_ , James Levinsohn, and Ariel Pakes**, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63* (4), 841–890.

**Cleanthous, Paris**, "Patient Welfare Implications of Innovation in the U.S. Antidepressant Market," 2002.

**Cockburn, Iain M., Jean O. Lanjouw, and Mark Schankerman**, "Patents and the global diffusion of new drugs," *American Economic Review*, 2016, *106* (1), 136–164.

**Dafny, Leemore S.**, "How do hospitals respond to price changes in emergency departments?," *American Economic Review*, 2005, *95* (5), 1525–1547.

**Danzon, Patricia M., Y. Richard Wang, and Liang Wang**, "The impact of price regulation on the launch delay of new drugs - Evidence from twenty-five major markets in the 1990s," *Health Economics*, 2005, *14* (3), 269–292.

**Dranove, D, C Garthwaite, and C Ody**, "How do nonprofits respond to negative wealth shocks? The impact of the 2008 stock market collapse on hospitals," *The RAND Journal of Economics*, 2017, *48* (2), 485–525.

**Dubois, Pierre and Laura Lasio**, "Identifying industry margins with price constraints: Structural estimation on pharmaceuticals," *American Economic Review*, 2018, *108* (12), 3685–3724.

**Duggan, Mark G.**, "Hospital ownership and public medical spending," *Quarterly Journal of Economics*, 2000, *115* (4), 1343–1373.

**Dunn, Abe**, "Drug innovations and welfare measures computed from market demand: The case of anti-cholesterol drugs," *American Economic Journal: Applied Economics*, 2012, *4* (3), 167–189.

**Duso, Tomaso, Annika Herr, and Moritz Suppliet**, "The welfare impact of parallel imports: A structural approach applied to the German market for oral anti-diabetics," *Health Economics*, 2014, *23*, 1036–1057.

**Evaluate Pharma**, "World Preview 2017, Outlook to 2022," Technical Report 2017.

**Gaynor, Martin and William B. Vogt**, "Competition among hospitals," *The RAND Journal of Economics*, 2003, *34* (4), 764–785.

**Grennan, Matthew and Ashley Swanson**, "Transparency and negotiated prices: The value of information in hospital-supplier bargaining," *Journal of Political Economy*, 2020, *128* (4), 1234–1268.

**Iizuka, Toshiaki**, "Experts' agency problems: Evidence from the prescription drug market in Japan," *The RAND Journal of Economics*, 2007, *38* (3), 844–862.

_ , "Physician Agency and Adoption of Generic Pharmaceutical," *American Economic Review*, 2012, *102* (6), 2826–2858.

**International Federation of Pharmaceutical Manufacturers & Associations**, "Biotheraputics Medicines," Technical Report 2012.

**Kaiser, Ulrich, Susan J. Mendez, Thomas Rønde, and Hannes Ullrich**, "Regulation of pharmaceutical prices: Evidence from a reference price reform in Denmark," *Journal of Health Economics*, 2014, *36* (1), 174–187.

# Appendix

Table A1: First Stage Estimation Results (The All-set Model)

| | |
|---|---|
| Price×Age | 0.000000219*** |
| | (6.60e-08) |
| Price×Age$^2$ | -7.67e-10 |
| | (5.86e-10) |
| Price×Sex | 0.00000120** |
| | (0.000000475) |
| Price×Bsa | 0.00000925*** |
| | (0.00000125) |
| Price×Pchemo | 0.00000753*** |
| | (0.000000961) |
| Drug×Age | 0.0393*** |
| | (0.00214) |
| Drug×Age$^2$ | -0.000485*** |
| | (0.0000198) |
| Drug×Sex | -0.0640*** |
| | (0.0171) |
| Drug×Bsa | -0.783*** |
| | (0.0442) |
| Drug×Pchemo | 2.588*** |
| | (0.0235) |
| Drug×D70 | 1.842*** |
| | (0.0383) |
| Filgrastim×D70 | 0.154*** |
| | (0.0350) |
| Nartograstim×D70 | 0.868*** |
| | (0.0637) |
| Lenograstim×D70 | 0.221*** |
| | (0.0367) |
| Generic×Age | 0.0240*** |
| | (0.00282) |
| Generic×Age$^2$ | -0.0000739*** |
| | (0.0000253) |
| Generic×Sex | -0.000405 |
| | (0.0209) |
| Generic×Bsa | 0.176*** |
| | (0.0555) |
| Generic×Pchemo | 0.203*** |
| | (0.0344) |
| ln(month + 1) | 0.936*** |
| | (0.0439) |
| product×year dummies | ✓ |
| $N$ | 184,954 |

Notes: We did not use a product choice set for each hospital, but used a common product choice set for all hospitals that included all products available at that time. "Price" indicates the wholesale price, calculated using Equation (1). Body surface area was calculated from the patients' weights and heights using the Du Bois Method (body surface area ($m^2$) = height(cm)$^{0.725}$ × weight(kg)$^{0.425}$ × 0.007184). "D70" is an indicator that takes the value of 1 if at least one of the disease codes (occasion of admission, comorbidity on admission, post-hospitalization onset disease, etc.) includes D70, and 0 otherwise. The robust standard errors are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively.

Table A2: Second Stage Estimation Results (The All-set Model)

| | (1) | (2) | (3) |
|---|---|---|---|
| | | product×year dummies | |
| | OLS | IV(Demand) | IV(Demand, Markup) |
| price | 0.000155*** | 0.000236** | 0.0000664 |
| | (0.0000440) | (0.000120) | (0.0000748) |
| product dummies | ✓ | ✓ | ✓ |
| $N$ | 136 | 136 | 136 |
| adj. R-squared | 0.928 | 0.925 | 0.925 |

Notes: We did not use a product choice set for each hospital but used a common product choice set for all hospitals that included all products available at that time. We used only those product $j$–year $t$ pairs for which $\delta_{jt}$ was precisely estimated where the z-values were greater than 2 in the first stage in Table A1. "Price" indicates the wholesale price, calculated using Equation (1). Column (1) shows the results of ordinary least squares regression. Column (2) shows the results of the instrumental variable regression, which uses "Demand IV" defined by Equation (15) and the interaction of "Demand IV" and a generic dummy as instruments. Column (3) shows the result of the instrumental variable regression, which also uses "Markup IV" defined by Equation (16) and the interaction of "Markup IV" and a generic dummy as instruments, in addition to the instruments in column (2). Standard errors are reported in parentheses. *, **, and *** denote 10%, 5%, and 1% significance levels, respectively.