

Addiction and Alcohol Tax: Evidence from Japanese Beer Industry *

Kohei Hayashida

University of California San Diego

Masakazu Ishihara

New York University

Makoto Mizuno

Meiji University

Kosuke Uetake

Yale School of Management

October 2023

Preliminary Please do not circulate without permission.

This paper studies the effects of taxation and regulation on addictive alcohol consumption. Exploiting the changes in tax policies and sales regulation in the Japanese beer market, we first show some descriptive evidence that consumers (i) are addicted to alcohol, (ii) are forward-looking and stockpile, but potentially present-biased, and (iii) substitute across categories in response to policy changes. To quantify the impacts of policy changes, we then estimate a dynamic structural model of alcohol purchase and consumption where consumers can be present-biased. A series of counterfactual simulations show that the current Japanese alcohol tax system is suboptimal in that alternative policies can increase tax revenues while keeping alcohol addiction lower. Finally, we derive the optimal alcohol tax policy, taking both externalities and internalities into account.

Key words: Addiction, Alcohol, Taxation, Time Inconsistency, Dynamic Structural Model

1. Introduction

Addiction to sin products such as drugs, tobacco, alcohol, and even soda, is a serious health issue in many countries. According to the National Survey on Drug Use and Health, 19.7 million American adults (aged 12 and older) suffered from a substance use disorder in 2017,

*First draft: January 18, 2023. This version: December 13, 2023. We thank Judy Chevalier, JP Dube, Yufeng Huang, Soheil Ghili, Julia Levine, Matthew Osborne, Helena Perrone, Jiwoong Shin, K Sudhir for their helpful comments. We also thank the seminar and conference participants at the Japan Institute of Marketing Science, the Marketing Dynamics Conference, the Marketing Science Conference, the Quantitative Marketing and Economics Conference, Yale. The paper's results are our own and do not represent the company's views.

and almost 74% of adults suffering from a substance use disorder in 2017 struggled with an alcohol use disorder. The same survey estimated that drug abuse and addiction cost American society more than \$740 billion annually due to direct costs such as healthcare expenses and crime-related costs and indirect costs such as less workplace productivity.

Policymakers regulate those markets in many ways from production to distribution. In Japan, for example, selling certain drugs such as cannabis is completely illegal. Even in the U.S., there are still many states that prohibit the recreational use of cannabis products. To control for excess consumption of addictive goods, it is more common that policymakers in many countries to impose taxes on those products. Most countries levy taxes on alcohol products and tobacco, and some states in the US have started imposing taxes on soda.¹

One of the rationales for such sin taxes is externalities. That is, people tend to under-appreciate the long-term consequences (e.g., addiction) such as health costs that consuming sin goods may have on themselves in the future. Hence, the tax may be able to carve out over-consumption by internalizing the future costs.

The taxes imposed on sin products (sin tax) are also important revenue sources for the government. Excise tax revenue from tobacco taxes totaled \$12.5 billion in 2019, accounting for 13 percent of all excise tax revenue, and revenue from alcoholic beverages amounted to \$10.0 billion in 2019, 10 percent of total excise receipts. Moreover, as of March 2022, U.S. states reported a combined total of \$11.2 billion in tax revenue from legalized recreational use of cannabis.

In this paper, we investigate the effects of taxes and regulations on the consumption of alcoholic beverages with a special emphasis on addiction. Our context is the Japanese beer markets where the alcohol taxes and regulations change over time.

The Japanese beer markets give us an ideal laboratory to investigate the research questions. First, there has been an ongoing discussion about whether the current alcohol tax system facilitates addiction by imposing a low tax rate on the category with higher alcohol content. In particular, doctors and addiction specialists in Japan have been raising alarms about some highball beverages with a 9-10% alcohol punch due to their addictiveness and the ease of inebriation they offer.² A similar concern arises in another country such as U.S.,

¹ For example, in January 2017, the city of Philadelphia introduced a “soda tax” on sweetened beverages, which is considered to be one of the major causes of obesity. Other cities and counties such as Berkeley (California) and Cook County (Illinois) also impose similar soda taxes to reduce the consumption of sugary beverages.

² For example, <https://toyokeizai.net/articles/-/336587?page=3> (in Japanese).

where canned cocktails or ready-to-eat drink beverages are charged by a higher alcohol tax than hard seltzer. Second, there have been and will be substantial policy changes in taxes and regulations on alcohol purchases. Among those, the alcohol-specific tax was updated in October 2020 to reduce the differences in tax rates across sub-categories of alcohol products. Also, the government is planning to unify the tax rates across categories by 2026. Besides the tax reforms, the government also prohibited excess sales of alcoholic beverages by large retail stores in June 2017 to protect small-scale liquor stores from competition.³ We exploit these policy changes as natural experiments to see how consumers' addiction behavior changes. Thus, it is highly policy relevant to study the impacts of alcohol taxes in the Japanese context and there are useful policy variations that allow us to credibly identify the effects of taxes on consumer behavior.

We begin with some descriptive analysis using the policy changes as natural experiments. First, we report that consumers do respond to policy changes. In particular, the category shares significantly change when the new alcohol tax policy is implemented in 2020. The consumer responses allow us to identify the intra-temporal substitution patterns across alcohol categories in response to policy changes. This also implies that taxation and regulation can maneuver addiction if they are carefully designed. Second, we show that consumers are indeed addicted to alcohol and forward-looking by using a series of reduced-form regression by [Becker et al. \(1994\)](#) and [Gruber and Köszegi \(2001\)](#). Our regression results based on [Becker et al. \(1994\)](#) suggest that consumers are rationally addicted to alcohol products as both the past consumption levels and the future price matter when they decide the current consumption level. In addition, we employ [Gruber and Köszegi \(2001\)](#)'s idea on using the exogenous tax shifts to identify the consumer's forward-looking behavior; If consumers are forward-looking and anticipate the future price increase, they may purchase more before the price hike or start to consume less to reduce their future consumption. Our results are consistent with the former one, that is consumers increase(decrease) their purchase amount in anticipation of the tax increase(decrease). Similarly, we show that consumers are forward-looking and hence adjust their stockpile of beverages in anticipation of future policy changes and sales, following the procedure by [Hendel and Nevo \(2006c\)](#). Lastly, we show some suggestive evidence of consumers' present bias. As in [Hinnosaar](#)

³ In addition, the general consumption tax (VAT) increased from 8% to 10% in October 2019.

(2016), a large fraction of consumers purchase alcohol products in two consecutive days. Moreover, when consumers purchase more than usual due to an exogenous shock (e.g., a policy change), they tend to consume and purchase alcohol products more frequently after the policy change.

Based on the findings from the descriptive analysis, we build a dynamic structural model of alcohol purchase and consumption. The structural model allows us to quantify the effects of existing alcohol tax and also investigate counterfactual tax policies. Our structural model extends the model in the previous papers such as [Gordon and Sun \(2015\)](#) and [Kim and Ishihara \(2021\)](#) by incorporating a few key features of our empirical context. First, we allow time inconsistency or present bias in the consumer's dynamic decision problem. It is important to consider such a behavioral bias since the implications of tax policies may significantly change as in [O'Donoghue and Rabin \(2006\)](#). Second, we model the degree of alcohol addiction to potentially depend on alcohol content. Although existing structural papers do not differentiate actual alcohol content from package size, it is well-known that the consumption of higher alcohol content is more likely to lead to addiction. Also, this feature allows us to examine counterfactual tax policies that depend on alcohol content.

We estimate the dynamic structural model by a Bayesian Markov chain Monte Carlo method proposed by [Imai et al. \(2009\)](#). We separately estimate the price process parameters so that consumers have beliefs about future prices. A major challenge of estimating our model with present-biased consumers beyond the previous papers such as [Gordon and Sun \(2015\)](#) and [Kim and Ishihara \(2021\)](#) is that there are two different types of value functions, i.e., the true value functions and the perceived value functions. Since consumers may have present-biased preferences, both actual and perceived value functions determine the consumers' actual choices over time (see, e.g., [Fang and Wang \(2015\)](#)).

The identification of the endogenous consumption model along with inventory is challenging because the researcher does not observe both consumption and inventory level in the data. We identify the model with an exclusion restriction that affects inventory costs, but not the flow utility (from purchase and consumption) such as electricity prices and prices of other items that take large space in a refrigerator. Also, the identification of the time inconsistency (or $\beta - \delta$ discounting), we exploit tax policy changes and their timing of announcement and enactment as in [Gruber and Köszegi \(2001\)](#) and [Chan \(2017\)](#). Since

the identification of discount factors is empirically challenging even if they are theoretically identified, we provide a series of Monte Carlo simulations (TBA).

Estimation results indicate some key findings. First, alcohol consumption is addictive, and people tend to consume more alcoholic beverages as they become more addicted. Second, consumers become more addicted when they have beverages that contain more alcohol given the same package size. Third, alcohol addiction depreciates as consumers stop drinking. Fourth, consumers are not perfectly rational, but they have present-bias preferences. Lastly, consumers' price elasticities significantly vary across categories, i.e., more expensive categories are less price elastic.

With the estimated structural model, we conduct counterfactual simulations to evaluate alternative alcohol tax designs. First, we consider the effects of the tax policy changes that the Japanese government is planning. This policy change aims for unifying different tax rates across categories into one single tax rate. The current plan is to implement this change in two steps: one in 2023, and another in 2026. We compare the purchase and consumption patterns under the planned tax reforms and those under the current tax system to see the effects of the policy changes on alcohol consumption. Second, we consider a hypothetical policy change where the government implements the policy change at once in 2020. Our preliminary simulation results show that (i) the planned change in alcohol taxes increases tax revenues for the government *and* reduces consumers' addiction level, (ii) the ahead-of-schedule policy change increases tax revenues relative to the planned change, while it decreases addiction. Therefore, the current alcohol tax system seems to facilitate addiction relative to alternative policies as suggested by Japanese doctors and addiction specialists.

We are planning to examine the role of the consumer's time-inconsistent preference for the effectiveness of tax policies. Also, we will examine an alternative tax policy where tax rates depend on alcohol content. Lastly, we are going to investigate the design of the optimal sin tax, taking externalities and internalities into account.

The rest of the paper is organized as follows. Section 2 describes the related literature and positions the current paper. Section 3 describes the institutional setting and data. Sections 5 and 6 describe the model and estimation respectively. Section 7 discusses the model estimates while Section 8 reports the findings from counterfactuals. Section 9 concludes.

2. Related Literature

This paper is related to several strands of the literature that study addictive goods' consumption and its regulations. The first strand of the related literature is the papers that empirically test the implications of the seminal theoretical paper by [Becker and Murphy \(1988\)](#) on rational addiction ([Becker et al. \(1994\)](#), [Grossman et al. \(1998\)](#), [Piccoli and Tiezzi \(2021\)](#), [Dragone and Raggi \(2021\)](#)). These papers derive the reduced-form models from the consumer's theory allowing addiction and estimate them. Roughly, the model describes the relationship between current consumption and past consumption, future consumption, past prices, future prices, and other variables. Our paper builds on these papers and shows some descriptive evidence of alcohol addiction before developing our dynamic structural model.⁴

The second strand of the related literature studies the dynamic structural models of addictive consumption/purchase ([Gordon and Sun \(2015\)](#), [Chen and Rao \(2020\)](#), [Kim and Ishihara \(2021\)](#)). These papers develop dynamic discrete choice models of alcohol purchases and consumption where consumers are "rationally" addictive as in [Becker and Murphy \(1988\)](#). Our paper adds to the literature by incorporating time-inconsistent preference into the dynamic model, which has potentially important implications for tax policy design. Also, our paper studies the effectiveness of the tax policies that are actually implemented in Japan using the structural model. In addition, our context contains much richer variations in the timing of the tax policy changes, which allow more precise identification of the discount factor and the time-inconsistent parameters.

The third literature that is related to ours is the papers that evaluate the existing tax changes or that examine optimal taxes, in particular, soda taxes (e.g., [Aguilar et al. \(2021\)](#), [Khan et al. \(2016\)](#), [Seiler et al. \(2021\)](#), [Bollinger and Sexton \(2018\)](#), [Griffith et al. \(2019\)](#), [Dubois et al. \(2020\)](#), and [Schmacker and Smed \(2023\)](#)), using the enactment of the tax as a natural experiment. Our paper differs from these papers in that we explicitly consider dynamic responses to tax policy changes by estimating a dynamic structural model of the category and amount choices, which allows us to investigate the long-run effects of counterfactual policy designs.

⁴ [Wang \(2015\)](#) highlight the importance of the storability of products to estimate price elasticity and the effectiveness of taxes by estimating a dynamic structural model of purchases. Her model does not take potential addiction into account.

Lastly, our paper is related to the literature on marketing / industrial organizations that study sin products (e.g., Conlon and Rao (2015), Conlon and Rao (2020), and Miller and Weinberg (2017), Hinnsaar (2016), and Hollenbeck and Uetake (2021)). Our paper complements those papers by providing additional evidence of the effects of sin taxation and regulation on consumer purchase behavior from a dynamic perspective.

3. Institutional Setting and Data

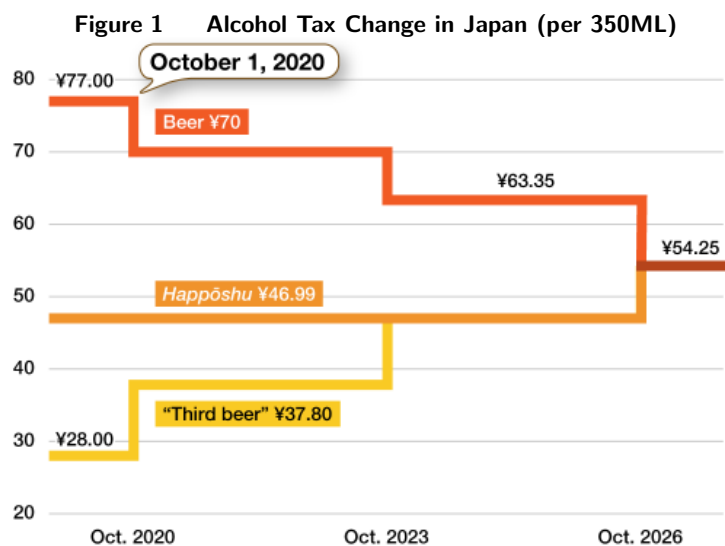
3.1. Japanese Beer Industry and Policy Reforms

In this section, we describe the general background of the Japanese beer industry and the regulatory schemes that we will exploit in our empirical analysis. In terms of absolute volumes, Japan is the seventh largest beer-consuming country, and per capita consumption is 40.1 liters of beer annually. The beer industry in Japan has been highly concentrated; The top-4 manufacturers (Asahi, Kirin, Suntory, and Sapporo) account for more than 99% of the market shares. Asahi and Kirin are two major manufacturers with more than 35% of market shares, respectively. Suntory and Sapporo follow the top two manufacturers with 15% and 11% of market shares.

An important policy related to alcohol consumption is alcohol taxation. The tax rate depends on the alcohol sub-categories. For beer, manufacturers produce three types of beers which consist of 1) *regular beer*, 2) *low-malt beer (Happoshu)*, and 3) *“the third” beer (Shin-genre)* in Japan. Although *low-malt beer* is made of the same ingredients as *regular beer*, it contains less malt. Hence, low-malt beers are typically cheaper than regular beer and have charged a lower alcohol tax rate as of 2019 (77.00 JPY/litter for beer and 46.99 JPY/litter for low-malt beer). *“The third” beer* tastes similar to regular beer, but it does not include any malt and its tax rate is even cheaper than low-malt beers (28 JPY/litter).⁵ The tax rate for other liquor such as cocktails and whisky, which we call the ready-to-drink alcoholic beverages (“RTD”) category, is the same as the third-beer category. The purpose of the alcohol tax reform that we will discuss in the following is to fix these disproportionate rates imposed on each category, and potentially increase the tax revenue.

In June 2017, the Japanese government announced updates on Liquor Tax Act and changes to alcohol tax for beer, law-malt beer, and third beer over 10 years. Figure 1 summarizes this alcohol tax change plan. In this series of updates, the taxation structure will

⁵ The third beer is technically not categorized as beer as it does not contain any malt, but is categorized as liquor under the current tax system. Hence, its tax rate is cheaper than low-malt beers. Historically, beer manufacturers introduced the third beer in response to an increase in alcohol tax for Happoshu (low-malt beer). In addition, the third beer is charged the lowest tax rate among the three types of beer.



Notes: Photo credit <https://www.nippon.com/en/japan-data/h00831/>

be simplified into three phases: October 2020, October 2023, and October 2026. In the first phase, which we will investigate in this paper, the alcohol tax for the beer category was reduced from 77 JPY/litter to 70 JPY/litter, while the tax for the third-beer category was increased from 28 JPY to 37.8 JPY. In this first phase, the tax for the law-malt category remains the same at 46.99 JPY. In 2026, all three categories will face the same tax level of 54.25 JPY. Due to this amendment, the tax-rate difference between the regular beer and the third beer is significantly reduced, which may affect substitution patterns between categories. In our empirical analysis, we will exploit these variations to see how alcohol taxation affects addictive consumption behavior. For the RTD category that includes cocktails, whiskey, and highball, the rate will be fixed at 28 JPY/litter till October 2026 and then raised to 35 JPY.⁶ Hence, the difference in tax rates between the beer category and the RTD category becomes smaller as well.

Furthermore, the Japanese government has been increasing the general consumption tax (VAT). Since first installed in 1989 at 3%, the consumption tax has been updated three times in 1997, 2014, and 2019. Our data includes before and after the last tax hike in October 2019, when the general consumption tax increased from 8% to 10%. In general, consumers do not see alcohol tax rates in price tags when they purchase alcoholic beverages, while they do see general sales taxes due to government regulation.

⁶ Alcohol tax rates for beer are higher than those for the RTD category because beer was historically categorized as a luxury good in Japan, and also the government aimed for protecting sake breweries.

In addition to the tax rate changes, in June 2017, the revised Liquor Tax Act restricts excessive sales of alcoholic beverages by large-scale retailers such as supermarket chains and discount stores. Although the Japanese Antitrust Law prohibits the dumping of alcoholic beverages, the revised Liquor Tax Act imposes further restrictions so that large-scale supermarkets and discount stores refrain from selling alcoholic beverages below their wholesale prices plus any related costs such as labor costs.⁷ The main purpose of this legislation is to protect small and mid-size liquor stores that may not receive quantity discounts from alcohol manufacturers. Another purpose is to increase tax revenues from alcohol purchases, though whether the regulation increases tax revenues or not is ambiguous.

We exploit these policy changes as natural experiments to identify consumer preferences including the extent that they are forward-looking and also present-biased. We will show some descriptive evidence on those points in the next section.

3.2. Data

We use the data from the Intage SCI/i-SSP on the Japanese beer markets from January 2014 to 31 March 2021. This data contains 1) the scanner panel data for alcohol products and non-alcohol beer products, 2) the household's TV advertising exposure data for the subset of the scanner panel monitors, and 3) the monitor's characteristics data (age, gender, family size, income, living prefecture, etc.).⁸

The scanner data of alcohol purchase tracks the household-level transaction across all the prefectures in Japan during the sample periods. The unit of observation in the scanner data is timestamp-ID (consumer ID)-SKU-level and includes information on the purchased quantity, paid price, and the visited store channel type such as supermarket, convenience store, drugstore, etc. We also obtained date-product level regular price information that is built based on the internal survey run by Intage. For the subset of the monitors, we have the household-level TV advertising exposure data. The unit of observation of the TV advertising data is date-ID-product-level. It contains information on the number of advertisements and the total minutes exposed during the day for each household. In the following, we discuss our way of constructing the estimation sample from these data.

⁷ In the case of violating the regulation, such retailers' name is published and their liquor license may be deprived.

⁸ The SCI panels are quota-sampled according to population proportions of age, gender, marital status, and region in Japanese residents aged 15-79. Among these monitors, i-SSP panels whose TV advertising exposures are measured are randomly collected by blocking the proportions of age, gender, and region. Therefore, the representativeness of the data used is guaranteed according to the company.

3.2.1. Estimation Sample Construction Since the original data contains many minor SKUs, we begin with selecting major products in our data for our analysis. To do so, we first aggregate SKUs into the same product’s brand level. The details of combining similar SKUs are in Appendix A.1. We use this product identifier to merge the scanner data with the advertising data below. To focus on the major products, we pick the top 62 sales products, which account for 66.7 percent of the entire sales in the raw data.

Then, we merge the scanner panel data with the TV advertising exposure data and the alcohol content data using the common product identifier. After joining them, there remain 6,334 households in 16 prefectures between 1st December 2014 and 31st March 2021.⁹ The final data is date-ID-product-route-level, and the total number of observations is 565,130.

Lastly, for each product, we collected alcohol percentage information from each manufacturer’s website. All the major brands showed up on the official websites.¹⁰ The details of this procedure are in Appendix A.2. In addition to product characteristics, our data contains information on the size (e.g., 350ML, 500ML) and the bundle pack size (e.g., 1 pack, 6 pack, 24 pack). Therefore, we can calculate the actual purchase volume for each purchase occasion, product, and consumer.

3.2.2. Summary Statistics Table 1 provides information on household demographics and consumer-level purchase records. According to the top panel of the table, the average age of the sample households is 45.7, and 70 percent of them are married couples. The median household income is 614 (10K JPY) per year (about 47,440 USD), and more than 80% of households own a car. Over half of the sample households come from the Tokyo metropolitan area (Tokyo, Kanagawa, Saitama, Chiba) and Osaka, although this information is not shown in the table.

The remaining panels of Table 1 present summary statistics of consumer-level purchase records. On average, consumers purchase a single category of alcoholic beverage in one purchase occasion, and the quantity purchased in terms of units of cans is 5.6, which suggests that consumers often buy pack items that contain multiple units of beer products. The average inter-purchase time is 16.4 days, and households visit at most 1.1 unique store formats per week, given that they visit any store at all.

⁹ Among them, 55 households did not show up in our advertising record possibly because they do not see advertising of the brands that we focused on.

¹⁰ We were originally concerned about the possibility that the product characteristics had changed and the information we saw at the time of the scraping is obsolete, especially for the older brands. Therefore, to double-check the

Table 1 Summary Statistics at the Household-level

Variable	Mean	Median	Std. Dev.	Min	Max
Demographics (ID-level)					
Income (1,000 JPY)	614.0	625	281.5	< 200.0	1,000 < .
Age	45.7	42	14.8	19.0	72 < .
Family Size	3.0	3	1.2	1.0	6 < .
Car Owned	0.8	1	0.4	0.0	1
Married	0.7	1	0.4	0.0	1
Purchase (ID-day-level)					
# of Category Purchased	1.1	1	0.4	1.0	5
Quantity Purchased (1 unit of can)	5.6	2	9.7	0.7	216
Expenditure for the Third-Beer (JPY)	275.8	0	799.6	0.0	28,350
Expenditure for the RTD (JPY)	81.9	0	244.6	0.0	11,526
Expenditure for Non-Alcohol (JPY)	33.8	0	218.2	0.0	9,910
Expenditure for Happoshu (JPY)	89.2	0	508.0	0.0	19,795
Expenditure for Beer (JPY)	257.9	0	887.0	0.0	34,128
Inter-Purchase Time (days)	16.4	5	52.3	1.0	2,156
Store visit (ID-week-level)					
# of Store Visited per week	1.1	1	0.4	1.0	5
Advertising exposure (ID-week-level)					
# of exposure per week	12.1	7	13.7	1.0	211

Note: The unit of observation to calculate the summary statistics is in the parenthesis of each header. The raw data of the demographic variables are discretized, thus we manually take the mid-point of each interval. For the bottom and top coding, we use the following criteria when calculating the mean, median, and SD: 1) we use 200 (10K JPY) as the lowest, and 1,000 as the highest for the income variable; 2) we use 20 as the minimum and the 72 as the maximum for the age variable; 3) we use 6 as the maximum family size. We convert pack products (e.g., 6 pack, 24 pack, etc.) into a can (e.g., 6 can, 24 can, etc.) when we calculate the Quantity Purchased.

Table 2 reports the summary statistics on different alcohol categories. The table includes summary statistics on product characteristics, such as alcohol content and price, for each category. Beer, Happoshu (low malt beer), and the Third Beer categories have an average alcohol content of around 5%, while the RTD category has the highest mean alcohol percentage. For prices, beer is the most expensive category, with an average price of about 201 JPY, while the Third Beer and RTD categories are cheaper. Finally, for market shares, the Third Beer category is the most purchased, followed by the RTD and the Beer categories.

Figure 2 illustrates the relationship between the mean price per unit and its alcohol percentage across different alcohol categories. Beer is comparatively expensive but has a moderate amount of alcohol content, while Third-beer is significantly cheaper than beer with similar alcohol content. Lastly, the products in the "RTD" category have the highest alcohol content among all the categories, despite being lower in price.

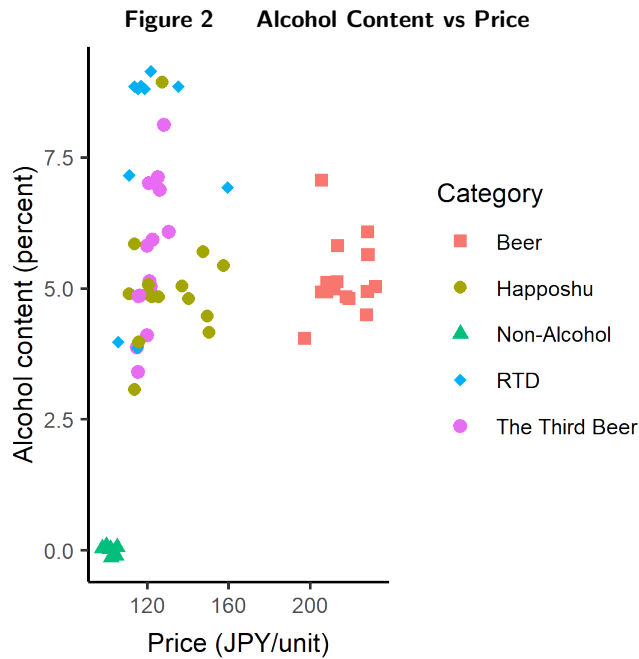
Figure 3 reports the shares of different consumer age groups for each alcohol percentage level. The figure indicates that younger individuals tend to buy alcohol with higher alcohol

consistency of the product characteristics over time, we used a few third-party databases, and Google image searches as well. Overall, we did not observe any inconsistencies in the alcohol content for the same product across time.

Table 2 Summary Statistics at the Year-Month-Product-level

Category	Mean	Median	Std. Dev.	Min	Max
Alcohol Content (pct) (Product-level)					
Beer	5.2	5.0	0.7	4.0	7.0
Happoshu	5.1	5.0	1.3	3.0	9.0
Non-Alcohol	0.0	0.0	0.0	0.0	0.0
RTD	7.4	9.0	2.1	4.0	9.0
The Third Beer	5.6	5.0	1.3	3.5	8.0
Price (JPY/unit) (year-month-product-level)					
Beer	201.0	201.4	3.7	192.1	207.6
Happoshu	132.4	132.1	4.2	125.4	141.3
Non-Alcohol	97.9	97.6	2.1	93.6	102.4
RTD	118.1	118.3	3.3	110.8	123.9
The Third Beer	108.3	107.2	3.6	103.2	119.5
Share (pct) (year-month-product-level)					
Beer	0.2	0.2	0.0	0.1	0.3
Happoshu	0.1	0.1	0.0	0.1	0.2
Non-Alcohol	0.1	0.1	0.0	0.0	0.1
RTD	0.1	0.1	0.0	0.1	0.2
The Third Beer	0.5	0.5	0.0	0.3	0.6

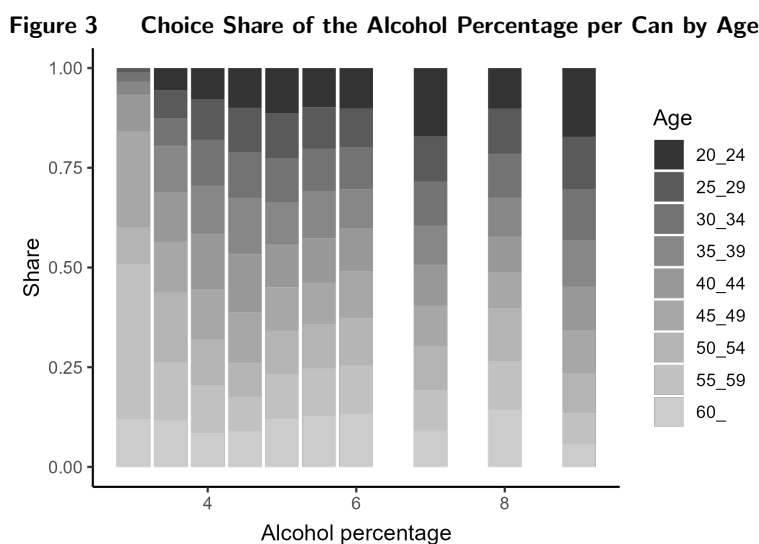
Note: The unit of observation to calculate the summary statistics is in the parenthesis of each header. We convert pack products (e.g., 6 pack, 24 pack, etc.) into a can (e.g., 6 can, 24 can, etc.) when we calculate the Quantity Purchased.



Notes: Each dot represents a product in our sample (62 in total). Price is calculated by taking the mean of the price in the ID-date-product level data.

content compared to other age groups. This, coupled with the lower prices and taxes on

these drinks, could be concerning as it may increase addiction among younger generations, which could persist for a long time.¹¹



Notes: To understand the share in the population as much as possible, we use the full sample consisting of 66,125 unique households instead of our estimation sample here. The x-axis is a per-can alcohol percentage. We first calculate the share of the purchased quantity of a specific alcohol percentage for each age class. Then, we calculate the share of the age class for each alcohol percentage. For example, the graph shows that the alcohol product with 9 percent alcohol content has the highest share among the age class 20-24.

3.3. Effect of Regulatory Reforms

In this section, we examine the effects of the regulatory reforms on prices, market shares, and consumer purchases.

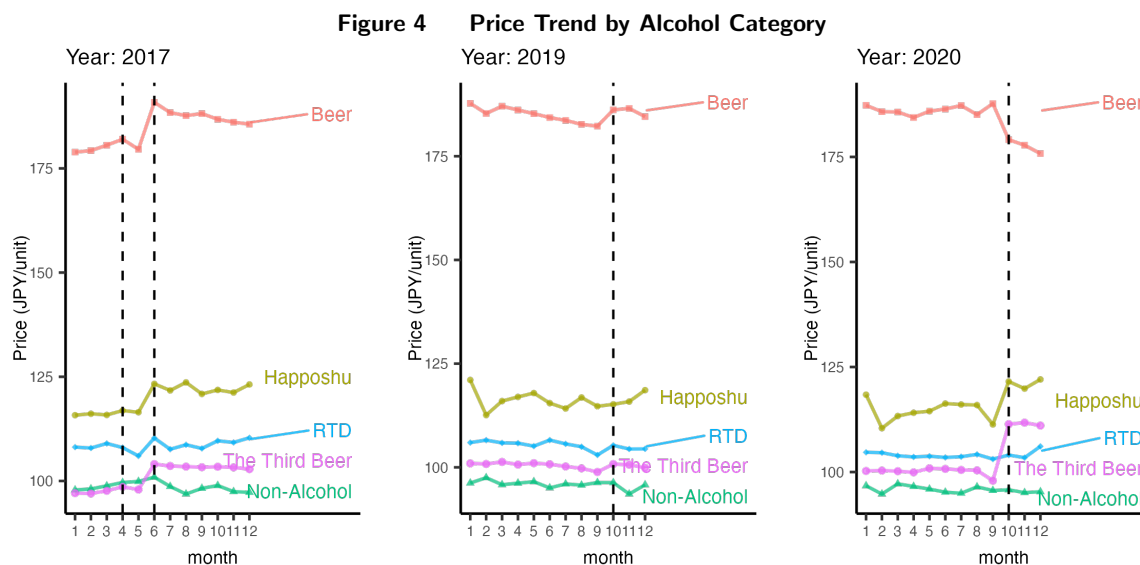
3.3.1. Prices by Category Figure 4 reports the (weighted) average price path by category around the policy change events. There are two policy-relevant events on the left panel; in April 2017, the Japanese government announced the alcohol tax policy will be restructured in 2020, 2023, and 2026. Also, in June 2017, the government published a guideline to prohibit excess discount sales of alcoholic beverages. The figure shows that prices

¹¹ According to the "Health Japan 21" national health promotion movement promoted by the Ministry of Health, Labour and Welfare, a "moderate and appropriate amount of alcohol consumption" is considered to be around 20 grams of pure alcohol per day on average. Also, they defined the "amount of alcohol consumption that increases the risk of lifestyle-related diseases" as a daily intake of 40 grams or more for men and 20 grams or more for women of pure alcohol in the "Health Japan 21 (Second Stage)" which started in 2013. Consuming 350ML of a can with 9 percent alcohol content amounts to around 25 mg of alcohol consumption, which exceeds the first criteria and the second criteria for women. The details can be found in <https://www.mhlw.go.jp/file/06-Seisakujouhou-10900000-Kenkoukyoku/0000047330.pdf>

did not respond to the announcement of the restructuring of the tax system, while prices increase significantly in all categories except the non-alcohol category when the discount regulation is put into place.

For the VAT hike (from 8% to 10%) in 2019 (middle panel), prices increase across categories, but the magnitude is smaller than the effects of other policy changes. Based on our calculations, the mean VAT pass-through rate to the retail prices is approximately 100%.¹² Also, prices do not change until the new general sales tax rate becomes effective.

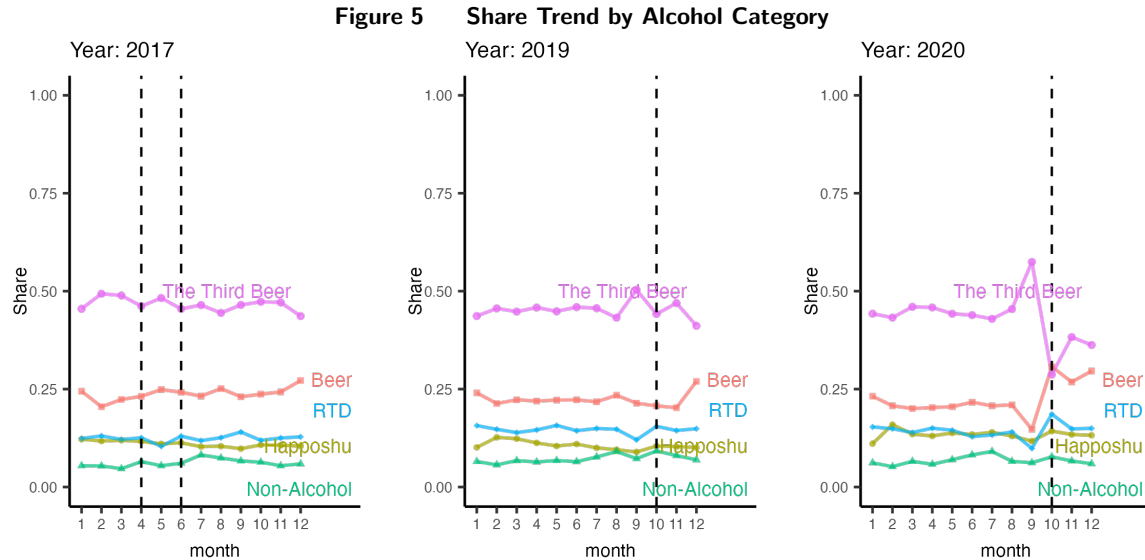
Lastly, for the change in the alcohol tax policy in 2020 (right panel), the average price of beer decreased, and the average price of the third beer increased significantly as expected. The average pass-through rate is also around 100% while there are variations by brands. An interesting observation is that the average price of Happoshu also increased even though the tax for the law-malt category was fixed at the same level. This would suggest the possibility of the substitution from the Third-beer to the Happoshu. As is also expected, we do not observe any price changes in the nonalcohol and RTD categories.



Notes: Prices are constructed as follows: 1) Each price is converted to per-can values. 2) The weighted mean of the per-can price is taken using the total units of cans in each cell (year-month-category) as a weight. The figures show only 350ML can products. The patterns are similar in 500ML can products as well. In April 2017, there was an announcement of Alcohol Tax Reform. In June 2017, there was a publication of the Guideline for Discount-Sale Regulation. In October 2019, there was a VAT increase from 8% to 10%. In October 2020, the Alcohol Tax Reform was enacted.

¹² To illustrate, let pre-tax price as P_0 and the original tax rate to be $t = 0.08$ and the post tax rate to be $t' = 0.10$. Then, before post-tax price $= (1 + t)P_0$ and after post-tax price $= (1 + t')P_0$. Therefore, $\frac{\text{after post-tax price}}{\text{before post-tax price}} = \frac{1+t'}{1+t} = 1.10/1.08 \approx 1.02$. So, if the pass-through is 100%, the relative price should increase around 2%. Our data indicate that the mean retail price increased by 2.3%.

3.3.2. Shares by Category Next, in Figure 5, we demonstrate the trend of category shares around the policy changes. Around the introduction of the regulation of discounts (left panel), market shares barely change. For the general tax hike (middle panel), the share of the Third beer category increases just before the general tax hike in October, which is likely due to a last-minute increase in demand. Due to the increase in Third Beer’s market share, RTD categories’ shares slightly decreased, indicating a close substitution between the two categories. The right panel shows that the market share of the Third beer category increases before the alcohol tax change due to the last-minute demand, and it significantly drops after the tax increase. By contrast, the beer category’s share drops before the tax change and increases after that. This also suggests the substitution between beer and the Third beer. Furthermore, the market share of beer and the Third beer categories after October 2020 looks different in levels compared to periods before.



Notes: The figures show the 350ML can products only. The patterns are similar in 500ML can products as well. In April 2017, there was an Announcement of Alcohol Tax Reform. In June 2017, there was a Publication of the Guideline for Discount-Sale Regulation. In October 2019, there was a VAT Increase from 8% to 10%. In October 2020, the Alcohol Tax Reform was enacted.

3.3.3. Alcohol Consumption Next, we investigate the effect of policy changes on alcohol consumption levels per household. To do so, we first run the following regression and estimate the residuals:

$$\text{Alcohol consumption}_{it} = \beta \text{Price}_{it} + \mu_i + \mu_t + \epsilon_{it}, \quad (1)$$

where $Price_{it}$ is a price index at prefecture r at time t , μ_i is household fixed effect, μ_t is year-month fixed effect, and ϵ_{it} is an iid error term.¹³

To deal with the endogeneity concern for $Price_{it}$, we instrument it by the exercise tax rate, the dummies for discount regulation in June 2017, and the dummies for alcohol tax policy change in October 2020. Our interest is how the residuals in the above regression explain the variations in alcohol consumption around the policy changes. To illustrate, we take the mean of the residuals ϵ_{it} across all the households and constructed the mean residuals $\epsilon_t \equiv \bar{\epsilon}_{it}$ across time.

Figure 6 shows how the mean residual changes over time. We find some large spikes right before the VAT reform in October 2019 and the alcohol tax policy changes in October 2020. There are a few key takeaways from this figure. First, the response to future policy changes seems to happen a month or at least a few months before the actual implementation. This may suggest that consumers take future price changes into account. We will look into and test this prediction more formally in the next section. Second, the alcohol consumption level seems to get smaller after the general sales tax hike (as the average residuals are almost below zero until the alcohol tax change in October).

4. Descriptive Analysis

In this section, we provide some descriptive evidence of addiction, stockpiling, and time inconsistency in our estimation sample to motivate our structural modeling.¹⁴

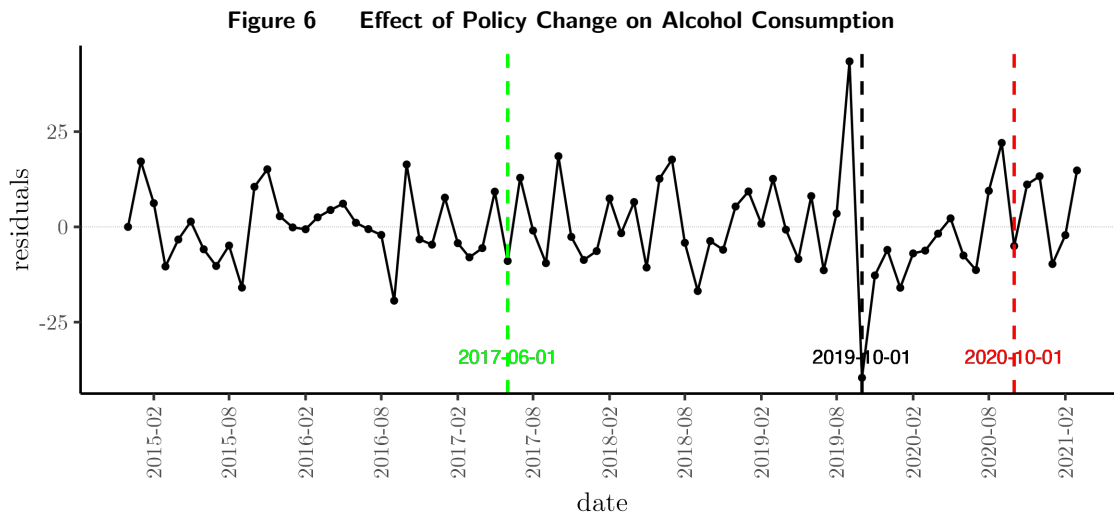
Note that both addiction and stockpiling lead to the correlation between past and current purchase decisions (Gordon and Sun 2015, Kim and Ishihara 2021). However, addiction leads to a positive correlation between past purchase decisions and current purchase decisions, while stockpiling generally creates a negative relationship.

4.1. Descriptive Evidence of Addiction and Forward-Looking Behavior

4.1.1. Is Alcohol Consumption Increasing based on the Past Consumption? We first show the evidence of addiction by comparing the within-household consumption changes

¹³ We calculated the alcohol consumption as follows: alcohol consumption (g) = $\frac{\text{alcohol content (percent)}}{100} \times \text{Size (ML)} \times \text{Specific gravity of alcohol}$. We use Specific gravity of alcohol = 0.8 according to one of the manufacturer's websites (https://www.suntory.co.jp/arp/proper_quantity/)

¹⁴ For this analysis, we focus on household-level panel observation instead of the product-level panel. Furthermore, we focus on households with sufficient observations. More precisely, we removed households if 1) the household's median inter-purchase days were longer than 31 days, or 2) if the households only had less than 10 times purchase records during the entire sample period. The selected sample contains 536,618 observations with 3,128 unique monitors.



Notes: The figure shows how the residual parts of the alcohol consumption regression change over time. The alcohol consumption regression uses household-month level data on alcohol consumption amount (mg), and we regress alcohol consumption amount on price index, the household fixed effects, and year-month fixed effects. The price index is constructed by taking the weighted mean of the price in scanner data for each prefecture, year, and month. We instrumented the Price index by the VAT rate, dummies for discount regulation in 2017/06, and alcohol tax policy changes in 2020/10. The green vertical dotted line indicates the discount regulation in 2017/06, the black vertical dotted line indicates the VAT reform in 2019/10, and the red vertical dotted line indicates the alcohol tax policy reform in 2020/10.

following [Gordon and Sun \(2015\)](#). In addictive products, consumption is expected to increase over time than decrease. To examine it, we calculate the frequency of an increase and a decrease in alcohol consumption for each household over time. That is, we calculate the empirical probability of increasing the purchase quantities of a consumer i as follows.

$$T_i^{-1} \sum_t^{T_i} I\{q_{i,t-1} < q_{i,t}\}, \quad (2)$$

where T_i is the number of purchase occasions and $I\{\cdot\}$ is an indicator function. We can similarly calculate the probability of remaining the same and decreasing, respectively.

In [Table 3](#), the first six columns are based on quantities of cans purchased while the last three columns are based on the alcohol content consumed. For quantity-based results, we calculate the probability of remaining the same, increasing, and decreasing for both alcoholic beverages and non-alcohol beverages. For consumption-based results, we calculate those probabilities based on consumed alcohol content aggregating all beverages. For each case, we show the results using all households (all), households who purchase alcoholic beverages above the median household (High), and those who are below the median household (Low).

Table 3 Testing Addiction

	Quantity-Based						Consumption-Based (mg)		
	Alcohol			Non-Alcohol			All	High	Low
	All	High	Low	All	High	Low			
Same	0.472	0.415	0.530	0.897	0.807	1	0.307	0.288	0.325
Increasing	0.266	0.295	0.237	0.052	0.097	0	0.349	0.359	0.340
Decreasing	0.262	0.290	0.233	0.051	0.095	0	0.344	0.353	0.335
t-stat	1.713	1.615	0.972	0.479	0.573	0	1.950	1.812	1.060
Std. error	0.003	0.003	0.004	0.002	0.003	0	0.003	0.003	0.004
N	3,128	1,565	1,563	3,128	1,678	1,450	3,128	1,564	1,564

Note: The unit of observation used to calculate each probability is ID-date-level. The High (Low) is from the households that have higher (lower) consumption than the household at the median. The t-stat shows the test statistic under the null hypothesis that increasing is equal to decreasing, and the alternative that increasing > decreasing.

The first three columns of Table 3 show that the increasing probability of purchase quantity is higher than the decreasing probability at a 5 percent level of statistical significance. If we focus on the subset of households that purchase alcohol above the median household, they have a higher probability of increasing alcohol quantity than decreasing it although it is not a statistically significant difference. Furthermore, we do not see statistically different patterns in non-alcohol products in columns 4-6. Additionally, when we do the same analysis for the alcohol consumption level instead of the purchased quantity, we again see a statistically significant difference at a 5 percent level, indicating the increasing probability is higher than that of decreasing especially in above-median-consumption households. Overall, the data suggest that alcohol products are addictive.

4.1.2. Regression-Based Rational Addiction Tests We further explore the existence of addiction by estimating the regression model based on the rational addiction model (Becker and Murphy 1988, Becker et al. 1994). Intuitively, if addiction exists, past consumption gives the higher marginal utility of current consumption. Also, the forward-looking consumer adjusts the current consumption based on future prices. We follow Dragone and Raggi (2021)’s AR1 specification to test the rational addiction to avoid the spurious addiction¹⁵:

$$C_{it} = \phi_0 + \phi_L C_{it-1} + \varphi_T P_{it} + \varphi_F P_{it+1} + v_i + d_t + u_{it}, \quad (3)$$

¹⁵ The canonical test of rational addiction regresses the current consumption on the lagged and future consumption, and the current price. Thus, it ends up with AR2 specification (Auld and Grootendorst 2004). Recently, Dragone and Raggi (2021) found that the canonical AR2 specification tends to predict spurious addiction such as in milk, and suggested using the AR1 specification.

where C_{it} is the individual i 's alcohol consumption (mg) in period t , P_{it} and P_{it+1} are the current price and one-period-ahead price, v_i is the individual fixed effects that capture time-invariant preferences that are correlated with lead and lagged consumption, d_t is time fixed effects, and u_{it} is an iid error term.

To estimate parameters consistently, it is necessary to deal with endogeneity problems on C_{it-1} , P_{it} , and P_{it+1} as these are correlated with unobserved error terms. To tackle this, we use the instrumental variables approach. The instruments for P_{it} , and P_{it+1} are beer and the third-beer's alcohol tax rate and consumption tax rate between $t-1$ and $t+2$. For C_{it-1} , we follow [Dragone and Raggi \(2021\)](#) and assumed it to be exogenous. We also include a dummy variable for alcohol tax change in October 2020, and a dummy variable for the sales regulation in June 2017. To control for the inter-temporal and household-level heterogeneity, we include year-month and household fixed effects.

Table 4 **Addiction Regression**

	Alcohol Consumption $_{it}$		
	Full (1)	Alcohol (mg-basis) (2)	Alcohol (ML) (3)
C_{it-1}	0.255 (0.006)	0.250 (0.006)	0.232 (0.005)
P_{it}	-2.20 (0.547)	-2.26 (0.547)	-2.29 (0.862)
P_{it+1}	1.31 (0.506)	1.23 (0.500)	1.76 (0.803)
R ²	0.48	0.46	0.41
Observations	145,154	141,149	141,149

Note: The unit of observation of the data used in the regression is monthly/household-level. The DV is $\log(\text{alcohol consumption})_{it}$. We included household FE, year/month FE, and prefecture FE in all the regression. For prices, we constructed the weighted mean as quantity sold as weights across products when we collapse the date/household/product-level dataset into monthly/household-level data.

Table 4 shows the estimated results using 2SLS.¹⁶ Column (1) shows the full result, showing that the lag variables in consumption and the lead variable in price are statistically

¹⁶ For this study, we aggregate the data at the ID-monthly level. When we aggregate into the ID-monthly level, we create a price index as follows. First, we convert pack prices into the per-unit-can price. Then, we take the weighted mean of the per-unit-can price with the total unit of cans as a weight for each year/month/prefecture-level observation.

significant. This suggests that the data support the view that consumers are forward-looking and addicted to alcohol.¹⁷ The coefficients of the current price φ_T are both negative and statistically significant as expected. Columns (2) and (3) show that this pattern is robust to the definition of consumption.

4.1.3. Addiction Test using Institutional Variations We provide additional evidence that households exhibit forward-looking behavior, building on [Gruber and Köszegi \(2001\)](#). Our test leverages institutional variations in tax rates to investigate whether consumers adjust their current consumption levels in anticipation of future changes. Since the alcohol tax rates change only for the beer category and the Third-Beer category, we focus on these two categories in this analysis. Of course, the consumption of other categories may change due to substitution.

Extending [Gruber and Köszegi \(2001\)](#)'s specification, we added our two tax policy changes as follows:

$$\begin{aligned} C_{ict} = & \alpha_1 \text{Effective VAT Rate}_t + \alpha_2 \text{Enacted VAT Rate}_t \\ & + \beta_1 \text{Effective Alcohol Tax Rate}_{ct} + \beta_2 \text{Enacted Alcohol Tax Rate}_{ct} \\ & + d_i + d_t + e_{it}, \end{aligned} \quad (4)$$

where C_{it} is the individual i 's alcohol consumption (mg) of category $c \in \{\text{Beer, Third-Beer}\}$ in period t , $\text{Effective VAT Rate}_{ct}$ and $\text{Effective Alcohol Tax Rate}_{ct}$ are the effective VAT tax rate in period t and the effective category c 's alcohol tax rate in period t , respectively, and $\text{Enacted Alcohol Tax Rate}_{ct}$ is category c 's alcohol tax rate already announced at period t , but not yet effective at that time.¹⁸ Thus, the value is as same as the effective rate before the announcement dates and after the effective dates. The model controls for the individual fixed effect, d_i , and the time fixed effect, d_t . Note that we run the regression separately by each category c .

We expect α_1 and β_1 to be negative as consumers dislike higher tax rates, while it is ambiguous whether β_2 is positive or negative. On the one hand, it can be positive to the

¹⁷ To see how we interpret the positive coefficient as addiction, see the following first-differenced equation.

$$\Delta C_{it} = \phi_L \Delta C_{it-1} + \varphi_T \Delta P_{it} + \varphi_F \Delta P_{it+1} + \Delta d_t + \Delta u_{it}.$$

Hence, the positive ϕ_L indicates that the consumption level increases.

¹⁸ The alcohol tax reform was announced in April 2017, which is almost 3 years before the actual implementation. Since it is unlikely that consumers were fully aware of or consider such a future policy change way before the actual implementation date, we vary the threshold that is perceived to be enacted.

extent that consumers increase stockpiling in anticipation of a future tax rate increase. On the other hand, it may be negative if addictive consumers reduce alcohol consumption now in order to avoid larger future expenses due to their increased addiction level and higher tax rates.

Table 5 **Addiction Regression (Gruber-Koszegi)**

	Alcohol Consumption _{ct}	
	Beer (1)	Third-Beer (2)
Effective VAT _t	-18.0 (3.88)	-47.8 (6.25)
Enacted VAT _t	12.3 (3.68)	42.8 (5.39)
Effective Alcohol Tax _{ct}	-3.63 (0.814)	-15.1 (1.55)
Enacted Alcohol Tax _{ct}	0.534 (0.714)	11.7 (1.34)
R ²	0.56	0.62
Observations	61,111	56,667

Note: The unit of observation used in the regression is monthly/household-level. The DV is Alcohol consumption_{it}. Enacted Alcohol Tax Rate_{ct} is category *c*'s alcohol tax rate announced but not yet effective at period *t*. Before the announcement and after the effective dates, the value is as same as the effective rate. Remember that beer's alcohol tax rate decreased while the third beer's tax rate increased in October 2020. We included household FE, and year/month FE in all the regressions. SEs are in parenthesis and clustered by households.

Table 5 shows the consumers' response to two tax policy changes. We find very strong negative effects of the current effective tax rates in both categories (the first and second rows). Moreover, we find different patterns in response to the expected change in the alcohol tax (the third row). For the Third Beer category where the alcohol tax increases, there is a strong positive increase, while for the Beer category where the alcohol tax decreases, there is a smaller insignificant impact. Hence, consumers seem to respond to a tax increase and a decrease in an asymmetric way.

Table 6 presents the sub-sample analysis across demographics. First, all segments responded to the future tax increase in the third-beer category while they are less responsive to the future price decrease in the beer category. Second, we find that households with a car are more responsive to the Third Beer category's future tax increase than those

without a car. This may be because households without a car find it harder to stockpile multiple units of packages earlier. Third, we did not find a systematic difference in income. Overall, our findings suggest that households plan ahead and adjust their consumption in anticipation of tax policy changes.

Table 6 Gruber-Koszegi Regression (Heterogeneity)

	Beer				Third Beer			
	Income		Car		Income		Car	
	Low	High	Yes	No	Low	High	Yes	No
Effective VAT _t	-18.3 (5.67)	-17.4 (6.14)	-19.7 (4.98)	-11.2 (5.45)	-40.0 (7.58)	-60.3 (11.5)	-48.8 (7.01)	-33.3 (12.6)
Enacted VAT _t	11.3 (4.80)	13.0 (6.43)	13.6 (4.63)	10.7 (5.59)	34.7 (6.73)	56.0 (10.0)	44.5 (6.16)	30.8 (11.3)
Effective Alcohol Tax _{ct}	-3.49 (1.02)	-4.00 (1.39)	-3.70 (1.01)	-3.68 (1.23)	-14.0 (1.86)	-16.7 (2.70)	-17.0 (1.91)	-8.19 (2.10)
Enacted Alcohol Tax _{ct}	0.639 (0.831)	0.574 (1.29)	0.686 (0.881)	0.505 (1.11)	11.0 (1.71)	13.0 (2.18)	13.4 (1.65)	6.13 (1.94)
R ²	0.52	0.61	0.58	0.51	0.61	0.65	0.61	0.68
Observations	28,580	25,475	44,817	12,993	29,484	21,238	43,163	10,862

Note: The unit of observation used in the regression is monthly/household-level. The DV is Alcohol consumption_{it}. Enacted Alcohol Tax Rate_{ct} is category *c*'s alcohol tax rate announced but not yet effective at period *t*. Before the announcement and after the effective dates, the value is as same as the effective rate. Remember that beer's alcohol tax rate decreased while the third beer's tax rate increased in October 2020. We included household FE, and year/month FE in all the regressions. SEs are in parenthesis and clustered by households.

4.2. Evidence on Stockpiling

We next test whether our sample households do stockpile for beer products following the procedure suggested in the literature (Gordon and Sun 2015, Kim and Ishihara 2021). Specifically, we test the following theoretical implications of the dynamic decision model derived in Hendel and Nevo (2006c): 1) Consumers buy more quantities when discounted and stockpile, 2) Consumers have a shorter inter-purchase time when discounted, 3) Consumers have a longer inter-purchase time after they bought when discounted. The first hypothesis simply mentions that price-sensitive consumers buy more during sales to stockpile them. The second hypothesis is supported if the consumer is willing to make a forward purchase to increase her inventory. Lastly, the third hypothesis is true if the increase in inventory reduces the consumer's need for the next purchase.

In Table 7, we examine the three hypotheses above by comparing the mean alcohol purchase and inter-purchase time with/without a sale and non-sale purchases.¹⁹ The first column shows the average during non-sale purchases. The second column (Sales diff, Total) displays the average during sale purchases minus the average during non-sale purchases, where we take an average across households and periods. In the third column (Sales diff, Within), we take the difference between each household's purchase during the sale and non-sale periods, and then take the average of those differences over households.²⁰

Overall, we find support for the prediction from the stockpiling behavior. First, the first row of Table 7 indicates that households purchase more during the sales periods, supporting the first hypothesis. Second, the column in the third row of Table 7 shows that the days since the last purchase is 0.486 days shorter when there is a discount, which is consistent with the second hypothesis. Note that the difference between the second column ("Total") and the third column ("Within") shows the importance of looking into the within-household variations rather than across-household variations as the latter may simply reflect the difference in the households' overall alcohol consumption tendency. Lastly, the fifth row of the table indicates that it takes longer until the next purchase if the household purchases alcohol during sales than during non-sale, which is consistent with the third hypothesis.

To show additional evidence that households do stockpile, we investigate whether households respond to tax increases by increasing the large pack share while decreasing the single pack share.²¹ Figure 7 shows the trend of the pack-level share by alcohol category. For example, in May 2017, there is a notable share transition in the Third-beer category from single pack to 6 or 24-pack (left panel). This is likely to be the response to the incoming price increase in June 2017. Similarly, in September 2019, consumers purchase larger packs

¹⁹ We define discount sales by measuring how far the purchased price is from the regular price. The regular price changes over time according to Intage's internal survey, thus we used its price corresponding to the specific period interval. We used the following procedure to determine whether each purchase occasion is a discount. First, we calculate the difference that purchase prices are below the regular price. Then, we define it as a discounted purchase occasion if the price difference is larger than some thresholds. In Table 7, we show the results using 30 percent as the threshold to define the sales. In this case, 29.2 percent of our observations are allocated to the discount purchase. The overall results are robust when we change the threshold to 10 or 20 percent. This procedure is similar in [Hitsch et al. \(2021\)](#).

²⁰ To calculate the difference controlling for the seasonality and region-specific factors, we estimate the mean difference in columns 2 and 3 in Table 7 by regressing outcome variable on intercept and fixed effects (monthly dummies and prefecture dummies).

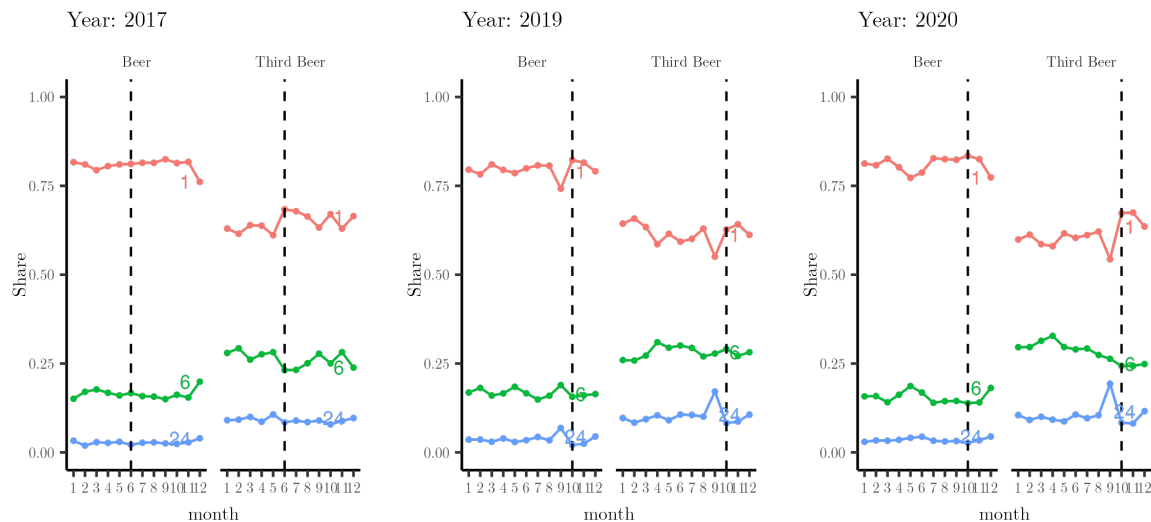
²¹ To ease the summary and analysis, we here limit the attention to the package size as either 1, 6, or 24 packs. This shares almost 99% of the purchase record in our sample data.

Table 7 Descriptive Analysis of Stockpiling

	Non-sale	Sale (diff)	
		Total	Within
Mean alcohol purchase (g)	75.198 (1.112)	72.032 (0.615)	49.706 (0.548)
Weeks from previous purchase	11.375 (0.273)	0.443 (0.101)	-0.486 (0.128)
Weeks until next purchase	11.163 (0.270)	0.918 (0.100)	0.210 (0.122)

Note: The Non-sales column reports the mean values for each row obtained from the observations restricted to non-sale purchases. Sales (diff): total column contains the difference in values averaged across all households. Sales (diff): within' column contains the average difference calculated within a household and then averaged across households. Robust standard errors are in parenthesis.

of beer, and the Third-beer categories prior to the price increase in October 2019 (middle panel). The same pattern applies in September 2020 for the Third-beer category as it is the category that price increased due to the Alcohol Tax reform while it was not the case for the beer category as expected (right panel). Thus, this alternative evidence also suggests that consumers rationally respond to exogenous price shocks by stockpiling prior to future prominent price changes.

Figure 7 Share Trend by Pack: Suggesting Stockpiling Behavior

Notes: The red represents 1-pack, the green represents 6-pack, and the blue represents 24-pack. In April 2017, there was an Announcement of Alcohol Tax Reform, and in June 2017, there was a Publication of the Guideline for Discount-Sale Regulation. In October 2019, there was a VAT Increase from 8% to 10%. In October 2020, the Alcohol Tax Reform was enacted.

Table 8 Fraction of Households with Consecutive Days Purchase of Alcohol Products

	Household-level			Purchase level
	at least once	more than twice	more than ten	
All	0.77	0.63	0.29	0.09
Household size				
1	0.74	0.59	0.26	0.10
2	0.72	0.57	0.26	0.09
3	0.67	0.52	0.22	0.08
4	0.68	0.53	0.23	0.08
5	0.64	0.50	0.21	0.08
6-	0.67	0.51	0.22	0.09
Car				
with car	0.76	0.61	0.28	0.09
without car	0.72	0.58	0.26	0.10
Category				
Beer	0.44	0.29	0.09	0.05
Third beer	0.49	0.35	0.12	0.06
Happoshu	0.30	0.18	0.06	0.05
RTD	0.46	0.32	0.12	0.06
Channel				
SM	0.59	0.45	0.17	0.06
CVS	0.46	0.31	0.12	0.07
Discount	0.16	0.07	0.02	0.02
Drugstore	0.19	0.10	0.03	0.02

Notes: The data we use to create this table is ID-date level with whether he/she did consecutive alcohol purchases (beer, third-beer, happoshu, RTD). In the first (second) column, we show the fraction of households that have at least one (two) consecutive purchases during the sample periods. In the fourth column, we take the mean of the within-household fraction of the consecutive purchases. That is, on average, 9 percent of their purchase is consecutive. In the bottom panel, SM refers to supermarkets and CVS to convenience stores.

4.3. Evidence on Time-Inconsistency

This section presents evidence that our sample households' behavior aligns with time inconsistency. To do so, we focus on each household's inter-purchase time (days). Although we do not directly observe consumption, inter-purchase time contains information on the consumption rate per household (Ailawadi and Neslin 1998).

4.3.1. Consecutive Alcohol Purchase Our first evidence of time inconsistency in alcohol purchases follows the idea in Hinno Saar (2016). Intuitively, time-inconsistent consumers tend to buy more frequently than otherwise. More specifically, if consumers buy alcoholic beverages on two consecutive days, they are likely to be time-inconsistent because they could have purchased multiple cans (with potentially a lower price) on the first day. Thus, consecutive purchases may reflect the degree of time inconsistency. In concrete, we calculate the fraction of households that have two consecutive days of alcohol purchase.

In Table 8, the first row shows that 77 percent of households did consecutive alcohol purchases at least once, 63 percent more than twice, and about 30 percent more than ten times. These fractions are greater than what is reported by [Hinnosaar \(2016\)](#) based on the US data.²² The fourth column shows the mean fraction of the consecutive visits across households to control for the entire purchase record length per household. This also shows the non-negligible fraction of purchases within a household is coming from consecutive ones.

Since people would purchase different kinds of alcoholic beverages on two consecutive days, in the middle panel of Table 8, we report the same statistics by category. If a household purchases the same category of products on two consecutive days, they are more likely to be time inconsistent. We find that about 44% of households who buy at least one beer product did a consecutive purchase of beer at least once, and about 10% of households did so more than ten times. The pattern is similar across categories.

Moreover, the bottom panel of Table 8 reports the same statistics by channel. Similar to the middle panel, if a household buys alcohol from the same channel on two consecutive days, then that provides further evidence of time inconsistency. Again, we find strong evidence of consecutive purchases of alcohol from the same channel, especially for supermarkets and convenience stores.

The alternative explanation for the consecutive alcohol purchases is addiction. A consumer who purchased yesterday and immediately consumed beer can be addicted due to it, and his/her utility of consuming another beer became higher. However, the rational consumer who has a reasonably high shopping trip cost should expect this consequence yesterday, thus he/she would have purchased more beer yesterday. In that case, the consumer does not have to purchase today but still is able to consume beer from his/her inventory. Hence, consumers who visit for two consecutive days may not be completely forward-looking due to some reasons other than addiction.

4.3.2. Inter-purchase Time Shrink before/after the Alcohol Tax Reform Next, we examine how the household's inter-purchase time responds to the alcohol tax reform. Note that the alcohol tax reform in 2020 exogenously increased the incentive for forward-looking

²² This may be partly because of the difference in shopping patterns: Japanese households do grocery shopping more frequently than the US ones. Even so, we believe there is a high fraction of households that do not plan ahead and purchase alcohol products day by day.

households to stockpile Third-beer products beforehand (i.e., September 2020), and delay purchasing beer products. If some households are time-inconsistent, those households that stockpile Third beer drink it at home faster than usual, and hence, the inter-purchase time for Third beer after the tax reform would be shorter than the inter-purchase time if they would not stockpile them. In contrast, for beer products, we expect that the inter-purchase time is not affected due to this tax reform as there was no incentive to further stockpile beer products.

To test whether households that stockpiled prior to October 2020 had a shorter inter-purchase time than usual controlling for potential confounds of households' inter-purchase time, we run the following regression:

$$\begin{aligned} \log(\text{interval}_{it}) = & \gamma_1 \text{Buy packs right before the ATR}_i + \gamma_2 \text{After ATR}_t \\ & + \gamma_3 \text{Buy packs right before the ATR}_i \times \text{After ATR}_t + \alpha \bar{c}_i \\ & + \text{Buy packs}_{it} + \beta \mathbf{X} + \epsilon_{it}, \end{aligned} \quad (5)$$

where interval_{it} is the inter-purchase time (days) for a household i at period t , After ATR_t is an indicator variable that is equal to 1 if the period t is after the Alcohol Tax Reform (ATR) in October 2020, Buy Packs_{it} is an indicator variable that is equal to 1 if a household i purchased at least one item of 6 or 24 pack at period t , $\text{Buy packs right before the ATR}_i$ is an indicator variable that is equal to one if a household i purchased at least one pack item of 6 or 24 pack at September 2020. \bar{c}_i is the average monthly alcohol consumption for a household i to control for the level of the inter-purchase time per household, \mathbf{X} are covariates including demographic variables such as age, gender, marriage status, number of household members, car ownership status, and income class.

Table 9 shows estimates for all and each category. We find that the households with more alcohol consumption in general purchase alcohol more frequently (and hence the inter-purchase time is smaller), and households purchase less frequently after buying packs as they stockpile more alcohol products. After the tax reform, the inter-purchase time for Third beer becomes longer as its prices increase, while that for the beer becomes shorter. Hence, overall, households purchase alcohol less frequently.

Our main interest is the fifth row, the estimate of the interaction term (γ_3). For the Third-beer, we find that the inter-purchase time becomes shorter for households that purchased packs right before the ATR. On the contrary, for Beer (column 3), it is not statistically

Table 9 Testing Time-Inconsistency by Inter-Purchase Interval

	All (1)	Third-beer (2)	Beer (3)	Happoshu (4)	Other (5)
Average Alcohol Consumption _{<i>i</i>}	-0.014 (0.000)	-0.012 (0.0001)	-0.014 (0.0002)	-0.013 (0.0004)	-0.019 (0.0003)
Buy Packs _{<i>it</i>}	0.458 (0.006)	0.484 (0.011)	0.374 (0.013)	0.450 (0.024)	0.523 (0.018)
After ATR _{<i>t</i>}	0.061 (0.014)	0.255 (0.032)	-0.057 (0.033)	0.091 (0.055)	0.054 (0.028)
Buy Packs right before the ATR _{<i>i</i>}	-0.013 (0.005)	-0.018 (0.011)	0.106 (0.013)	-0.189 (0.031)	0.174 (0.013)
After ATR _{<i>t</i>} × Buy Packs right before the ATR _{<i>i</i>}	-0.081 (0.014)	-0.087 (0.029)	-0.043 (0.033)	-0.126 (0.051)	-0.035 (0.032)
R ²	0.16	0.15	0.17	0.24	0.17
Observations	280,744	82,304	65,214	23,082	68,578

Note: The dependent variable is $\log(\text{inter-purchase-days}_{it})$. ATR is the abbreviation for the Alcohol Tax Reform. The unit of observation used in regression is household-date-level. For this analysis, we used the subset of the households that have purchase records both before and after the Alcohol Tax Reform on 1st October 2020 (1,676 unique households). Standard errors are heteroscedasticity robust. Year-Monthly FE and prefecture FE are included. The coefficients of demographics, average consumption and *Buy packs_{it}* are omitted due to the space limit.

different from zero as expected. Therefore, the households increase the consumption rate more than they usually do after they stock more. This finding is consistent with time-inconsistency.²³

Although the regression results provide suggestive evidence of time inconsistency, another potential explanation for such results is simply due to addiction. That is, households who stockpile more than usual drink more as they become more addictive, which leads to a shorter inter-purchase time after the policy change.

To see this, we split the sample in terms of the average alcohol consumption and estimate the same models with the data on households who consume fewer alcohol products. Specifically, we define households with less alcohol addiction if their mean purchase amount is below the 75 percentile among all the households. Since those households are less likely to be addicted to alcohol, their Third-beer consumption before and after the alcohol tax reform may not be driven by the addiction.

Table 10 shows that is true. We find that households with less addiction still behave in a time-inconsistent manner. We also find that the effect is statistically significant in

²³ Other covariates also have reasonable signs: 1) the high average alcohol consumption households have shorter inter-purchase times, 2) the inter-purchase time is usually long after purchasing bulk packages, 3) owning a car increases the inter-purchase time. Some covariates show interesting insights into this market: 1) larger family has a shorter inter-purchase time, 2) income effect on the inter-purchase time is non-monotonic, 3) younger people have a shorter inter-purchase time, etc. The results are robust to the specification of the model too. We got almost similar interpretations when we apply the Cox-proportional Hazard model instead of the linear regression.

the Third-beer category, suggesting that households who stockpiled in anticipation of the future price increase have shorter inter-purchase time even if we focus on less addicted households. Therefore, addiction only cannot explain the consumers' behavior, and we need to take consumers' time inconsistency into account.

Table 10 Subsample Analysis: Below 75 percentile Alcohol Consumption Households

	All (1)	Third-beer (2)	Beer (3)	Happoshu (4)	Other (5)
Buy Packs _{it}	0.425 (0.006)	0.422 (0.013)	0.353 (0.016)	0.266 (0.028)	0.500 (0.019)
After ATR _t	0.036 (0.015)	0.240 (0.035)	-0.106 (0.036)	0.110 (0.059)	0.041 (0.027)
Buy Packs right before the ATR _i	0.046 (0.006)	0.115 (0.013)	0.138 (0.016)	-0.053 (0.039)	0.185 (0.013)
Average Alcohol Consumption _i	-0.024 (0.0001)	-0.021 (0.0002)	-0.022 (0.0005)	-0.024 (0.0008)	-0.025 (0.0002)
After ATR _t × Buy Packs right before the ATR _i	-0.047 (0.015)	-0.063 (0.033)	-0.012 (0.038)	-0.154 (0.056)	-0.018 (0.032)
R ²	0.22	0.22	0.20	0.30	0.21
Observations	216,797	55,960	49,386	16,304	67,350

Note: The dependent variable is $\log(interval_{it})$. ATR is the abbreviation for the Alcohol Tax Reform. The unit of observation used in regression is household-date-level. We only used the households that have purchase records both before and after the Alcohol Tax Reform on 1st October 2020. It resulted in 1,676 unique households. Standard errors are heteroscedasticity robust. Year-Monthly FE and prefecture FE are included. The coefficients of some demographics are omitted for the space limit but included in the regression.

To investigate the robustness of this result, we also run a similar test around the VAT reform in 2019 October. Since this policy increased the price in all the categories, we only show the aggregated result in Table 9. Similar to the previous result, we find the households with less addiction still shrunk the inter-purchase time when they stockpiled due to the policy. Thus, these results suggest that our sample households have time-inconsistency in alcohol product purchase, so we develop the structural model that allows the time time-inconsistency in the next section.

5. Structural Model

This section proposes a model of forward-looking consumers' alcohol drink purchase/consumption decisions. Our model extends [Gordon and Sun \(2015\)](#) and [Kim and Ishihara \(2021\)](#) by incorporating a few key features of our empirical context. First, while those existing papers do not study the observed policy changes, we incorporate observed tax changes during our sample period into our model. We assume that consumers know the

Table 11 VAT Reform: Heterogeneity by Alcohol Consumption Level

	Below 25 percentile (1)	Below Median (2)	Below 75 percentile (3)
Buy Packs _{it}	0.442 (0.024)	0.432 (0.011)	0.425 (0.006)
After ATR _t	0.007 (0.025)	0.038 (0.017)	0.016 (0.014)
Buy Packs right before the ATR _i	-0.040 (0.019)	0.002 (0.009)	0.083 (0.006)
Average Alcohol Consumption _i	-0.060 (0.0004)	-0.047 (0.0002)	-0.025 (0.0001)
After ATR _t × Buy Packs right before the ATR _i	-0.141 (0.029)	-0.109 (0.016)	-0.064 (0.010)
R ²	0.35	0.27	0.22
Observations	65,543	150,755	236,787

Note: The dependent variable is $\log(\text{interval}_{it})$. VATR is the abbreviation for the VAT Reform in October 2019. The unit of observation used in regression is household-date-level. We only used the households that have purchase records both before and after the VAT Reform. Standard errors are heteroscedasticity robust. Year-Monthly FE and prefecture FE are included. The coefficients of some demographics are omitted for the space limit but included in the regression.

timing of the tax change, and thus the time to a tax policy change is one of the state variables in our model. This variable plays a key role in identifying consumers' time preferences. Second, we allow for the possibility that consumers have time-inconsistent preferences. It is important to consider time inconsistency for the purchase and consumption decisions of addictive goods as consumers may behave differently in the long run. Finally, we allow the extent that consumers are addicted to alcohol to be dependent on the total ethanol content as suggested by the medical literature. Hence, a consumer is more likely to be addicted when she drinks one 300ml cocktail with 5% alcohol per volume than when she drinks one 750ml beer with 1% alcohol per volume.

5.1. Single-Period Utility

We suppose that at each time ($t = 1, 2, \dots, \infty$), consumers $i = 1, 2, \dots, I$ visit a store exogenously with probability ϕ_i . Upon visiting a store, consumers choose which product category to buy ($j = 1, \dots, J$) as well as how much to buy (from a discrete set of quantities q). Consumers can also choose not to buy any alcohol products ($j = 0, q = 0$). We let $d_{ijqt} = 1$ if consumer i buys quantity q of category j at time t , with $\sum_{jq} d_{ijqt} = 1$. After the purchase stage, consumers choose how much to consume. The available inventory is given by the starting inventory level I_{it} at the beginning of period t plus the quantity purchased (q). If consumers do not visit a store at time t , they just decide how much to consume based on I_{it} .

Now we explain key state variables. First, we let a_{it} be the level of addiction for consumer i at time t , p_t is a vector of *unit* prices at time t ($p_t = \{p_{jqt}\}_{jq}$), and $\tau_t \geq 0$ be the time to tax implementation ($\tau_t = 0$ implies the taxes are in effect).²⁴ Then, consumer i 's single-period utility from purchase decision $d_{it} = \{d_{ijqt}\}_{jq}$ and consumption decision c_{it} in state $s_{it} = (a_{it}, I_{it}, p_t, \tau_t)$ is

$$u(d_{it}, c_{it}, s_{it}, \epsilon_{it}; \theta_i) = u_c(c_{it}, a_{it}; \theta_i) + u_p(d_{it}, p_t, \tau_t, \epsilon_{it}; \theta_i) + H(c_{it}, d_{it}, I_{it}; \theta_i), \quad (6)$$

where $u_c(c_{it}, a_{it}; \theta_i)$ is the consumption utility, $u_p(d_{it}, p_t, \tau_t, \epsilon_{it}; \theta_i)$ is the purchase utility, and $H(c_{it}, d_{it}, I_{it}; \theta_i)$ is the inventory holding cost. ϵ_{it} in u_p is an idiosyncratic error to the purchase utility, and θ_i is a vector of parameters. We model the consumption utility as

$$u_c(c_{it}, a_{it}; \theta_i) = \omega_{i1}c_{it} + \omega_{i2}c_{it}^2 + \omega_{i3}a_{it} + \omega_{i4}a_{it}^2 + \omega_{i5}a_{it}c_{it}. \quad (7)$$

The consumption utility includes both linear and quadratic effects of consumption level c_{it} to capture the decreasing marginal utility from consumption (if $\omega_{i2} < 0$). This utility incorporates the reinforcement and tolerance effects of rationally addicted behavior: ω_{i3} and ω_{i4} capture the overall impact of addiction level on consumption utility. Specifically, when $\omega_{i4} < 0$, a high addiction level will decrease the overall consumption utility (tolerance effect); $\omega_{i5} > 0$ implies that the marginal utility from consumption increases in addiction level, and captures the reinforcement effect. Moreover, we assume that consumers do not distinguish categories at the consumption stage (e.g., [Erdem et al. 2003](#)). However, we will account for the difference in alcohol contents across categories in the evolution of addiction levels.

The consumer's addiction level evolves as

$$a_{it+1} = (1 - \kappa)a_{it} + c_{it}^a, \quad (8)$$

where the consumer's addiction in the next period depends on the current addiction level, a_{it} , alcohol-content adjusted consumption, c_{it}^a , and the depreciation rate of addiction level ($\kappa \in [0, 1]$). We introduce alcohol-content-adjusted consumption because we expect that a higher alcohol content leads to a higher addiction level. Our assumption here is that when consumers consume alcohol products, alcohol contents do not influence their consumption

²⁴ We assume $\tau_{t+1} = \tau_t - 1$ if $\tau_t > 0$, and $\tau_{t+1} = 0$ if $\tau_t = 0$.

amount and consumption utility (all that matters is volume). However, the addiction level is influenced by the amount of alcohol content in the consumption. We will define c_{it}^a below.

We model consumer i 's purchase utility as

$$u_p(d_{it}, p_t, \epsilon_{it}; \theta_i) = \sum_{jq} d_{ijqt} \left(\alpha_i (p_{jq} + \text{tax}_j \mathbb{I}(\tau_t = 0)) \times q + \gamma_i \ln(\text{adv}_{ijt}) + \xi_{ijq} + \epsilon_{ijqt} \right), \quad (9)$$

where $p_{jq} \times q$ is the total price, and tax_j is the unit tax imposed on category j (which can be positive or negative, as in Figure 1), which is applicable only when $\tau_t = 0$. We denote category j 's TV advertising exposure counts for consumer i at time t by adv_{ijt} .²⁵ We then let α_i be the price sensitivity parameter, γ_i be the sensitivity to advertising, ξ_{ijq} capture consumer heterogeneity in category and quantity preferences, and ϵ_{ijqt} is an idiosyncratic error.²⁶ We note that any quantities that are not consumed at time t will be stored in inventory for future consumption. Thus, we assume that the inventory level evolves as

$$I_{it+1} = I_{it} + \sum_{jq} d_{ijqt} \times q - c_{it}. \quad (10)$$

To capture the influence of alcohol contents on the evolution of addiction level, we introduce an alcohol-content adjusted inventory, which evolves as

$$I_{it+1}^a = I_{it}^a + \sum_{jq} d_{ijqt} \times q \times \mu_j - c_{it}^a. \quad (11)$$

where μ_j is the adjustment of alcohol contents for category j , $\mu_j = \exp(\rho AC_j)$, where AC_j is the average alcohol content of products in category j and ρ is a parameter to be estimated. Now, the alcohol-adjusted consumption (c_{it}^a) is then defined as

$$c_{it}^a = c_{it} \frac{I_{it}^a + \sum_{jq} d_{ijqt} q \mu_j}{I_{it} + \sum_{jq} d_{ijqt} q}.$$

if $c_{it} > 0$, and $c_{it}^a = 0$ otherwise. Here, similar to the quality-weighted inventory in [Erdem et al. \(2003\)](#), we assume that categories are consumed proportionally so we can define the alcohol-content adjusted consumption using the ratio of the two types of available inventory.

²⁵ We assume that consumers view TV advertising as a random shock and do not anticipate future TV advertising in the forward-looking decision.

²⁶ In our empirical application, we assume that ξ_{ijq} is additively separable, i.e., $\xi_{ijq} = \xi_{ij} + \xi_{iq}$.

A key variable to identify the present-biased preference in our model is the time to tax implementation (τ_t). This state variable transitions deterministically as follows:

$$\tau_t = \begin{cases} \tau_{t-1} - 1 & \text{if } t \notin \mathbf{T} \\ T_{l+1} - T_l & \text{if } t = T_l, \end{cases} \quad (12)$$

where $\mathbf{T} = \{T_l\}$ is an ordered set of periods with policy changes. We assume that consumers perfectly know those policy changes happen. Since the government makes public announcements about important policies way before the actual implementations, this is a reasonable assumption.

Finally, the consumer's inventory holding cost is based on the average inventory level during the period, i.e., the average of the beginning inventory after purchase ($I_{it} + \sum_{jq} d_{ijqt}q$) and the ending inventory (i.e., I_{it+1}), which simplifies to

$$H(c_{it}, d_{it}, I_{it}; \theta_i) = (\eta_{i0} + \eta_{i1} Z_{it}) \left(I_{it} + \sum_{jq} d_{ijqt}q - \frac{c_{it}}{2} \right), \quad (13)$$

where (η_{i0}, η_{i1}) measures the unit inventory holding cost for consumer i , and Z_{it} includes cost shifters for the holding cost such as electricity prices and other input costs. Note that these cost shifters affect only inventory holding costs, but not consumption nor purchase utilities. Consumers cannot consume more than current inventory ($c_{it} \leq I_{it} + \sum_{jq} d_{ijqt}q$) at time t . We assume that all categories exert equal utility at the moment of consumption after alcohol adjustments (Erdem et al. 2003; Gordon and Sun 2015; Hendel and Nevo 2006a; Kwon et al. 2021; Wang 2015). This assumption lessens the computational burden as we do not need to keep track of the inventory for each category and thus it decreases the size of the state space significantly.²⁷

5.2. Price Expectations

Following Erdem et al. (2003), we use p_{jqt} to denote the unit price of category j and quantity q time t , and the price follows a first-order Markov process. To capture the volume discount, we divide the grid points for purchase quantities into two volume levels: high volumes and low volumes. We then assume that the unit price for high volumes is a function of the unit price for low volumes so that consumers only form future price expectations for the unit price for low volumes.

²⁷ We refer the reader to Hendel and Nevo (2006b) for detailed justifications of the assumption.

Specifically, we assume that

$$\ln p_{jt}^L = \psi_{0j}^L + \psi_{1j}^L \ln p_{j,t-1}^L + \psi_{2j}^L \ln \bar{p}_{-j,t-1}^L + \epsilon_{jt}^L, \quad (14)$$

$$\ln p_{jt}^H = \psi_{0j}^H + \psi_{1j}^H \ln p_{jt}^H + \epsilon_{jt}^H, \quad (15)$$

where we assume that the vector, $\{\epsilon_{jt}^L\}_j$, follows a multivariate normal distribution with zero mean and covariance Σ^H , and that ϵ_{jt}^L follows an independent normal distribution with zero mean and standard deviation σ_j^H . We estimate the price process parameters jointly with the demand-side parameters in a Bayesian framework.

5.3. Dynamic Optimization Problem

At time t , consumers not only care about the current utility from the current purchase and consumption but also the future utilities. Let δ be the (long-run) discount factor, β be the present bias parameter, and ψ be a vector of all price process parameters. As in the existing papers (e.g., Fang and Silverman (2009), O'Donoghue and Rabin (2006)), an agent is solving an infinite horizon dynamic optimization problem by thinking of the individual as consisting of many autonomous selves, one for each period. In each period a time t self maximizes her current utility, assuming that her future selves control their subsequent decisions. Below we derive the dynamic behavior of a *sophisticated* consumer. We define that an agent is sophisticated if the agent in period t correctly knows her future selves' present bias and anticipates their behavior when making her period- t decision.

Let $(\sigma_{it}^V, \sigma_{it}^W)$ be a strategy profile when consumer i visits a store at time t (thus makes both purchase and consumption decisions) and when consumer i does not visit a store at all (thus makes only consumption decision), respectively. A strategy profile specifies her action in all possible states and under all possible realizations of unobserved shock. Also, denote the continuation strategy profile from period t on as $\sigma_{it}^{V+} \equiv \{\sigma_{ik}^{V+}\}_{k=t}^{\infty}$ and $\sigma_{it}^{W+} \equiv \{\sigma_{ik}^{W+}\}_{k=t}^{\infty}$. To describe the equilibrium of the intra-personal game of an agent, we introduce two different types of the value functions. First, let V and W be the agent's period- t expected continuation utility for a given continuation strategy profile $(\sigma_{it}^{V+}, \sigma_{it}^{W+})$ when the consumer visits a store and when she does not visit any store, respectively. These value functions can be thought of as hypothetical value functions when her own present bias is eliminated.

$$V(s_{it}, \epsilon_{it}; \theta_i, \phi_i, \psi; \sigma_{it}^{V+}, \sigma_{it}^{W+}) = u(s_{it}, \epsilon_{it}; \theta_i; \sigma_{it}^V)$$

$$\begin{aligned}
& + \delta E_{s_{it+1}} \left[\phi_i E_{\epsilon_{it+1}} V(s_{it+1}, \epsilon_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{V+}, \sigma_{it+1}^{W+}) \right. \\
& \left. + (1 - \phi_i) W(s_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{V+}, \sigma_{it+1}^{W+}) | s_{it}, \sigma_{it}^V \right] \quad (16)
\end{aligned}$$

$$\begin{aligned}
W(s_{it}; \theta_i, \phi_i, \psi; \sigma_{it}^{V+}, \sigma_{it}^{W+}) &= u_c(a_{it}; \theta_i; \sigma_{it}^W) + H(I_{it}; \theta_i; \sigma_{it}^W) \\
& + \delta E_{s_{it+1}} \left[\phi_i E_{\epsilon_{it+1}} V(s_{it+1}, \epsilon_{it}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{V+}, \sigma_{it+1}^{W+}) \right. \\
& \left. + (1 - \phi_i) W(s_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{V+}, \sigma_{it+1}^{W+}) | s_{it}, \sigma_{it}^W \right] \quad (17)
\end{aligned}$$

We can then define a perception-perfect strategy profile for a sophisticated consumer as a strategy profile $\sigma_i^* \equiv \{\sigma_{it}^{*V}, \sigma_{it}^{*W}\}_{t=1}^{\infty}$ such that

$$\begin{aligned}
\sigma_{it}^{*V}(s_{it}, \epsilon_{it}) &= \arg \max_{c_{it}, d_{it}} \left\{ u(c_{it}, d_{it}, s_{it}, \epsilon_{it}; \theta_i) \right. \\
& + \beta \delta E_{s_{it+1}} \left[\phi_i E_{\epsilon_{it+1}} V(s_{it+1}, \epsilon_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) \right. \\
& \left. \left. + (1 - \phi_i) W(s_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) | s_{it}, c_{it}, d_{it} \right] \right\}, \quad (18)
\end{aligned}$$

$$\begin{aligned}
\sigma_{it}^{*W}(s_{it}) &= \arg \max_{c_{it}} \left\{ u_c(c_{it}, a_{it}; \theta_i) + H(c_{it}, d_{it} = 0, I_{it}; \theta_i) \right. \\
& + \beta \delta E_{s_{it+1}} \left[\phi_i E_{\epsilon_{it+1}} V(s_{it+1}, \epsilon_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) \right. \\
& \left. \left. + (1 - \phi_i) W(s_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) | s_{it}, c_{it}, d_{it} = 0 \right] \right\}. \quad (19)
\end{aligned}$$

Here, the continuation payoff is discounted by $\beta\delta$ instead of δ as the sophisticated agent knows that the future selves are present-biased. Following the derivation in [Fang and Wang \(2015\)](#), we can derive the updating rule for $V(s_{it})$ and $W(s_{it})$ and solve them to get the choice-specific value functions for constructing the choice probabilities. Let us define the expected continuation payoff under a perception-perfect strategy profile as

$$Y(s_{it}; \theta_i, \phi_i, \psi; \sigma_{it}^{*V+}, \sigma_{it}^{*W+}) \equiv \phi_i E_{\epsilon_{it}} V(s_{it}, \epsilon_{it}; \theta_i, \phi_i, \psi; \sigma_{it}^{*V+}, \sigma_{it}^{*W+}) + (1 - \phi_i) W(s_{it}; \theta_i, \phi_i, \psi; \sigma_{it}^{*V+}, \sigma_{it}^{*W+})$$

and the current choice-specific value function as

$$\begin{aligned}
W_{c_{it}, d_{it}}^s(s_{it}, \epsilon_{it}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) &= u(c_{it}, d_{it}, s_{it}, \epsilon_{it}; \theta_i) \\
& + \beta \delta E_{s_{it+1}} [Y(s_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) | c_{it}, d_{it}, s_{it}]
\end{aligned}$$

$$\begin{aligned}
W_{c_{it}}^s(s_{it}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) &= u_c(c_{it}, a_{it}; \theta_i) + H(c_{it}, d_{it} = 0, I_{it}; \theta_i) \\
& + \beta \delta E_{s_{it+1}} [Y(s_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) | c_{it}, s_{it}]
\end{aligned}$$

Using the definition of $(V, W, V_{c,d}^s, W_c^s)$, we can write the choice-specific value functions as a function of only (V, W) as

$$\begin{aligned} V_{c_{it}, d_{it}}(s_{it}, \epsilon_{it}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) &= V_{c_{it}, d_{it}}^s(s_{it}, \epsilon_{it}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) \\ &\quad + (1 - \beta)\delta E_{s_{it+1}}[Y(s_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) | c_{it}, d_{it}, s_{it}] \\ W_{c_{it}}(s_{it}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) &= W_{c_{it}}^s(s_{it}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) \\ &\quad + (1 - \beta)\delta E_{s_{it+1}}[Y(s_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) | c_{it}, s_{it}] \end{aligned}$$

Now $W_{c_{it}}$ is not subject to an idiosyncratic error, we have

$$W(s_{it}; \theta_i, \phi_i, \psi; \sigma_{it}^{*V+}, \sigma_{it}^{*W+}) = \max_{c_{it}} W_{c_{it}}(s_{it}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+})$$

We can obtain V as

$$\begin{aligned} E_{\epsilon_{it}} V(s_{it}, \epsilon_{it}; \theta_i, \phi_i, \psi; \sigma_{it}^{*V+}, \sigma_{it}^{*W+}) &= E_{\epsilon_{it}} \max_{c_{it}, d_{it}} V_{c_{it}, d_{it}}(s_{it}, \epsilon_{it}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) \\ &= E_{\epsilon_{it}} \max_{c_{it}, d_{it}} V_{c_{it}, d_{it}}^s(s_{it}, \epsilon_{it}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) \\ &\quad + (1 - \beta)\delta \sum_{c_{it}, d_{it}} \left\{ \Pr(c_{it}, d_{it}) E_{s_{it+1}}[Y(s_{it+1}; \theta_i, \phi_i, \psi; \sigma_{it+1}^{*V+}, \sigma_{it+1}^{*W+}) | c_{it}, d_{it}, s_{it}] \right\}. \end{aligned}$$

We can obtain V, W by using these two recursive equations.

6. Estimation and Identification

6.1. Estimation

For the estimation of the proposed structural model, we use households who purchased alcohol products at least 20 times during our sample period. Also, since our sample period is long (over 7 years), we drop households whose maximum interpurchase time is greater than half a year and whose average interpurchase time is greater than two months. Moreover, we drop households whose maximum purchase volume per week, category, and store type is very large (greater than 8400 ml). Our final estimation sample consists of 1005 households.

We estimate the proposed discrete choice dynamic programming model using the Bayesian Markov chain Monte Carlo algorithm proposed by [Imai et al. \(2009\)](#). The model has three sets of parameters: (1) preference parameters $(\theta_i, \beta, \delta)$, (2) store visit parameters (ϕ_i) , and (3) price process parameters (ψ) . For preference parameters that are heterogeneous across consumers (e.g., ω_i), we assume that the population distribution follows an

independent normal distribution. We assume diffuse normal and inverted Gamma priors on the mean and standard deviation of the population distribution, so that the posterior distributions are normal and inverted Gamma, respectively. Then, we draw individual-level parameters using the independent Metropolis-Hastings algorithm (the proposal distribution is set to the population distribution), and draw the population mean and standard deviation parameters directly from the corresponding posterior distributions (i.e., normal and inverted Gamma distributions, respectively). For preference parameters that are homogeneous across consumers (e.g., ρ), we assume flat prior and draw them using the random-walk Metropolis-Hastings algorithm.

Store visit parameters are assumed to be heterogeneous. Since it follows a multinomial distribution (no visit, visit store type r), we assume a Dirichlet prior and draw directly from the posterior distribution, which is Dirichlet. For price process parameters, linear parameters are assumed to have a diffuse normal prior and drawn directly from the posterior distribution (which is normal) conditional on variance-covariance parameters. Variance-covariance parameters for the low volume unit price are assumed to have diffuse inverted Wishart prior and drawn directly from the posterior distribution (which is inverted Wishart) conditional on the linear parameters. For the high volume unit price, variance parameters are assumed to have diffuse inverted Gamma prior and drawn directly from the posterior distribution (which is inverted Gamma) conditional on the linear parameter.²⁸

In our application, we discretize quantity decisions into 13 most frequently purchased quantities (in 350ml). Observed quantities that do not fall into these 13 sizes are approximated by the nearest grid point. The value function computation requires the optimization of consumption. We discretized possible consumption amount using 6 grid points. The state variables of the proposed model include addiction (a_{it}), inventory (I_{it}), alcohol-adjusted inventory (I_{it}^a), prices (p_t), and the time to tax implementation (τ_t). The first four variables are continuous, so we use random-grids to compute value functions. Since addiction, inventory, and alcohol-adjusted inventory evolve deterministically conditional on purchase and consumption choices, we use the local interpolation approach (Imai et al. (2009)). For the low volume unit price, we use the price expectation processes to integrate out future prices.

²⁸ We note that while the price expectation process parameters enter the value functions, we do not use the information from purchases to infer the parameters of the price expectation processes.

For the high volume unit price, we use a set of independent normal draws to integrate it out conditional on a high volume unit price.

Regarding the timing of the tax changes, in the proposed model in Section 5, we have only one τ_t (time to tax implementation). But in our empirical application, we use five regulation and tax changes to liquor prices and taxes: (1) discount regulation (6/2017), (2) sales tax increase (10/2019), (3) liquor tax change 1 (10/2020), (4) liquor tax change 2 (10/2023), and (5) liquor tax change 3 (10/2026). For estimation, we have data up to only 3/2021, so we need to consider the first three changes and thus we need to have three τ_t 's (i.e., time counter for each exogenous event). In counterfactual experiments, we also need to consider the last two changes. Consumers learn about each exogenous event when it is formally announced, so in principle, τ_t tracks the time since the announcement to implementation. However, announcements are typically made well in advance and consumers hardly change their behavior until the time to the event gets closer. So we assume that consumers will start incorporating each exogenous event into their continuation value four weeks prior to its actual implementation. To capture each event, we need to compute four types of value functions: (1) value functions when no changes are expected, (2) value functions when consumers expect that the discount regulation will be implemented ($\tau_t^1 > 0$) and when it is already in effect ($\tau_t^1 = 0$) but consumers are unaware that the sales tax will be implemented, (3) value functions when the discount regulation is in effect (τ_t^1) and consumers expect that the sale tax change will be implemented ($\tau_t^2 > 0$) and when it is already in effect ($\tau_t^2 = 0$) but consumers are unaware that the liquor tax will be implemented, and (4) value functions when both discount regulation and sales tax change are in effect ($\tau_t^1 = \tau_t^2 = 0$), and consumers expect that the liquor tax will be implemented ($\tau_t^3 > 0$) and when it is already in effect ($\tau_t^3 = 0$) but consumers are unaware that further liquor tax changes will be implemented. Let $k = 1, 2, 3, 4$ index the four types of value functions. We can then summarize the value functions as (V_h^k, W^k) . We solve the value functions and estimate the model using the IJC algorithm, i.e., in each MCMC iteration, we compute and store pseudo-value functions for each of the above value functions at each state s .

We deal with the initial condition problem for $(a_{it}, I_{it}, I_{it}^a)$ as follows. We set these values to zero at time $t = 1$ (i.e., when a consumer made the first alcohol product purchase in the entire data set). Then, we use the next 12 weeks' data (observed prices and purchases) and

the model to simulate the values of these state variables at $t = 13$. The likelihood computation uses data from $t = 13$ only. Finally, in our current model, consumers visit at most one store type and purchase at most one category. However, in the data, in rare occasions, consumers visit multiple store types and/or purchase multiple categories. To correctly account for inventory and addiction, we need to account for all those purchases. The estimation handles this situation as follows. For the likelihood computation, we assume that consumers made those decisions independently. That is, when consumers visited multiple store types, we assume that each visit was independent (i.e., assuming that the other visit didn't happen), derive the likelihood for each occasion (i.e., it is possible that consumers visited both store types but did not buy anything), and count all likelihood contributions. When consumers purchased multiple categories (from one store type or both), we again assume that each purchase was independent, and count all the likelihood contributions. Then, for updating the state variables, we pool all the purchases (across stores or across categories), let consumers maximize the consumption given the inventory and all the purchases, and derive the next period state variables $(a_{it+1}, I_{it+1}, I_{it+1}^a)$. This way, we can reflect all the purchases on inventory and addiction evolution.

6.2. Identification

We now informally discuss the identification of our model. Key identification challenges come from i) endogenous consumption, and ii) time-inconsistent discount factor.

The identification of the model with endogenous consumption and stockpiling is challenging since the researcher does not have data on consumption and also inventory levels. Hence, the consumption utility and the inventory costs are usually not separately identified without strong assumptions on consumption (e.g., constant consumption). The identification requires an exclusion restriction that affects the inventory cost but not the purchase or consumption utilities. As examples of such exclusion restrictions, one can consider electricity prices, house rent, or milk prices. Those variables shift the cost of storing alcohol in the refrigerator but are unlikely to directly affect consumption or purchase utility. Intuitively speaking, if electricity prices go to infinity, then inventory costs go to infinity exogenously. Then, consumers do not stockpile at all, and hence consumption is equal to the purchase amount. Hence, we can observe consumption, and only dynamics can arise from addiction. In the appendix, we show some indirect evidence that those excluded variables are correlated with purchases.

Now we can identify consumption and purchase utilities. Addiction parameters can be identified by the variation in alcohol contents conditional on the same amount. They face the same inventory cost, while different consumption utilities. The joint distribution of purchase propensity and quantity choices identifies the coefficients on the consumption amount (c_{it}) in the consumption utility. Intuitively, if consumers purchase the same amount at every shopping trip, a shorter inter-purchase time indicates a higher consumption rate or if consumers purchase at the same time intervals, a larger purchase indicates a higher consumption rate.

Next, for the parameters in the purchasing utility, the inter-temporal price variations such as sales and consumers' responses identify the price coefficient as in the previous papers (e.g., [Hendel and Nevo \(2006a\)](#)). Similarly, the inter-temporal variations in advertising and consumers' responses allow us to identify the advertising coefficient. Lastly, consumers' variations in product category choices and quantity choices help to identify the fixed effects for purchase options. Once those two functions are identified, the inventory costs can be identified from purchase frequencies or purchase hazards in [Ching and Osborne \(2020\)](#).

A key departure of our model from the literature is the possibility of consumers' present-biased preference. As in Section 5, we model time inconsistency by incorporating $\beta - \delta$ discount factors (e.g., [Laibson \(1997\)](#), [Fang and Silverman \(2009\)](#)) and assume that consumers have sophisticated present-bias.²⁹ As well known in the literature, identifying the discount factor (without time inconsistency) in the dynamic model is challenging. As [Magnac and Thesmar \(2002\)](#) and [Abbring and Daljord \(2020\)](#) and [Abbring et al. \(2018\)](#) show, exclusion restrictions can be used to identify the discount factors. In particular, it is useful to use exclusion variables that affect the transition probabilities of states over time but do not affect the current utility. With variations in excluded variables, one can compare consumers who face different inter-temporal trade-offs conditional on the same current utility level. With time-inconsistency, it is even more challenging to separately identify two types of discount factors, β and δ . A few recent papers tackle this challenge including [Wang et al. \(2022\)](#), [Tsubota \(2021\)](#), [Abbring and Daljord \(2020\)](#) show that some exclusion restrictions are important for identification. Following this spirit, [Chan \(2017\)](#) estimates a model

²⁹ The observational data we have does not allow us to identify the degree of consumer naivete. Some papers estimate the degree of naivete using the demand for commitment devices in randomized experiments (e.g., [Bai et al. \(2021\)](#), [Miller et al. \(2022\)](#)).

with time-inconsistent consumer who decides to stay in a welfare program, and considers the number of remaining periods of welfare eligibility under the time limit to identify the present-biased preference.

Even if it is theoretically possible to separately identify β and δ with exclusion restrictions, [Abbring and Daljord \(2020\)](#)'s monte carlo simulations illustrate the identification of β and δ is empirically difficult because the Bellman equation typically contains the product of β and δ . The identification comes from the fact that the Bellman equation includes $\beta\delta$ or $\beta\delta^2$, i.e., the non-linearity in these two parameters provides identification. Empirically, what matters is the extent to which exclusion restrictions create sufficient variations in continuous payoffs so that β and δ are separately identified.

In our setup, we can take advantage of multiple policy changes in the data. In particular, time till a policy change can work as an exclusion restriction. Time till a policy change does not affect the current utility function (conditional on observables) as such a policy is still not effective, while it affects the consumer's belief and transition probabilities. This idea of identification is similar to [Gruber and Köszegi \(2001\)](#) and [Chan \(2017\)](#).³⁰ Moreover, as [Kong et al. \(2022\)](#) discuss, the dynamic discrete choice model of differentiated products has a built-in exclusion restriction that the flow payoff depends only on chosen alternative's characteristics but not on other alternatives' characteristics. This creates additional restrictions to help us identify both β and δ . That said, the identification is still empirically challenging and our identification relies on parametric assumptions. To support our identification argument, we conduct a series of Monte Carlo simulations and show that the parameters are reasonably identified with the variations we have.

7. Results

7.1. Parameter Estimates

We simulate 10,000 draws from the posterior distribution of the parameters, and use every 10th draw after 5,000 burn-in period to make the statistical inference. [Table 12](#) reports the population mean and population standard deviations of the parameters in the structural model. We begin with the parameters in the consumption utility. First, the parameters for consumption utility show a decreasing marginal utility ($\bar{\omega}_1 > 0, \bar{\omega}_2 < 0$). These estimates translates to the (static) optimal consumption of about three 350 ml cans per week when

³⁰ In particular, [Chan \(2017\)](#) considers the number of remaining periods of welfare eligibility under the time limit to identify the present-biased preference.

Table 12 Parameter Estimates

Parameter	Description	Population mean		Population sd	
		mean	sd	mean	sd
ω_1	consumption (linear)	0.510**	0.012	0.244**	0.008
ω_2	consumption (quadratic)	-0.078**	0.003	0.062**	0.001
ω_3	addiction (linear)	1.011**	0.024	0.546**	0.014
ω_4	addiction (quadratic)	-0.185**	0.003	0.072**	0.002
ω_5	consumption x addiction	0.264**	0.003	0.082**	0.002
η	inventory cost	-1.317**	0.022	0.712**	0.020
α	price sensitivity	-0.006**	0.001	0.033**	0.001
γ	advertising sensitivity	0.202**	0.013	0.225**	0.006
κ	addiction depreciation	0.900**	0.005		
ρ	alcohol effect on addiction	0.036**	0.0002		
β	present bias	0.801**	0.073		
δ	discount factor	0.983**	0.011		

Notes: ** indicates zero is not contained in the 1% credible interval. The population distributions of ξ_{ijq} and ζ_{ih} are estimated but not reported due to space constraint.

the addition level is zero. Second, parameters for addiction utility also indicate a decreasing marginal utility ($\bar{\omega}_1 > 0, \bar{\omega}_2 < 0$), and when the addiction level gets too large, additional addiction reduces the consumption utility. For example, when consumers do not consume alcohol products, the addiction starts lowering the utility at around 2.75. The interaction effect of consumption and addiction (ω_5) supports the reinforcement effect. For example, as the addiction level increases from zero to one, the above optimal consumption increases from three 350 ml cans per week to five 350 ml cans per week. Lastly, the inventory cost is negative as we expected.

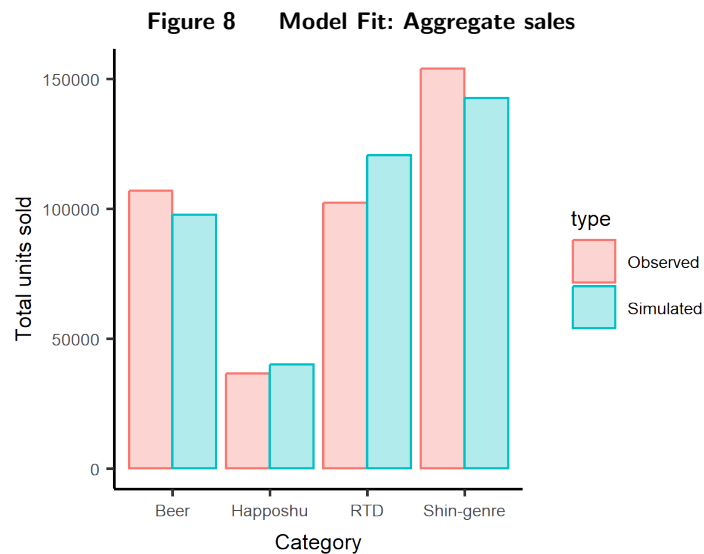
We then discuss the parameters in the purchase utility. The price sensitivity is negative and statistically significant. Note that the magnitude is relatively small because we use the Japanese yen as the price (which is usually about 100 times bigger numbers than the U.S. dollar). We also find that the advertising sensitivity is positive.

For the parameters that affect dynamic choices, the addiction depreciation rate is estimated to be 0.9, which is large. The estimate implies that the addiction level becomes half in about 8 weeks. Moreover, the alcohol effect on addiction is positive (0.036), suggesting that the consumption of a category with a higher alcohol content will make consumers more addicted.

The literature finds that the discount factor is heterogeneous across people. [Chan \(2017\)](#), for example, finds that around half of the individuals have a discount factor lower than 0.9. Also, in [Fang and Silverman \(2009\)](#), the sophisticated present-biased model yields a

discount factor of 0.88. Ferrall (2012) is the first to estimate a structural model with a distribution of discount factors and finds that 37% of the individuals have a discount factor of 1. For the present-bias parameter, the previous paper finds it even more heterogeneous. The present bias factor in Chan (2017) has a mean of 0.59, while Fang and Wang (2015) estimate that the present bias factor is 0.68.

Finally, we show the model fit (aggregate sales across consumers and weeks, and sales over time) in Figures 8 and 9. It shows that our proposed model is able to explain data well. Note that the periodic drops in sales in Figure 9 are due to the drop in store visits on new year's Eve.



Notes: Each bar indicates the total quantity sold per category from the model and data in our estimation sample. When aggregating, we count the unit of cans irrespective of the product's size such as 350ML, or 500ML.

7.2. Price Elasticities

We compute the price elasticities and quantify the substitution patterns among the four categories based on the structural parameter estimates. Table presents the results. The top panel uses the entire sample period, and the next two panels show the results of the pre-and post-liquor tax change. Each number shows the percentage change in demand for the column category in response to a 1% change in the price of the row category. Overall, the elasticities are consistent with the previous literature (e.g., Ishihara et al. (2022)). We have a few important observations. First, as we expected, the price elasticities for Third



Notes: Each point indicates the aggregated quantity sold in a week for each category from the model and data in our estimation sample. When aggregating, we count the unit of cans irrespective of the product's size such as 350ML, or 500ML. The dips in each panel are due to the low store traffic on new years' eve. Weeks in the x-axis are normalized to start from zero.

Table 13 Price Elasticities

	Beer	Happoshu	Third-beer	RTD
<i>All periods</i>				
Beer	-1.708	0.175	0.214	0.125
Happoshu	0.065	-2.091	0.163	0.071
Third-beer	0.203	0.418	-4.172	0.512
RTD	0.108	0.166	0.465	-3.943
<i>Pre-liquor tax change</i>				
Beer	-1.730	0.179	0.209	0.126
Happoshu	0.062	-2.209	0.144	0.072
Third-beer	0.204	0.405	-4.229	0.518
RTD	0.106	0.176	0.453	-4.035
<i>Post-liquor tax change</i>				
Beer	-1.483	0.148	0.291	0.117
Happoshu	0.090	-1.377	0.493	0.060
Third-beer	0.196	0.499	-3.200	0.470
RTD	0.126	0.103	0.673	-3.187

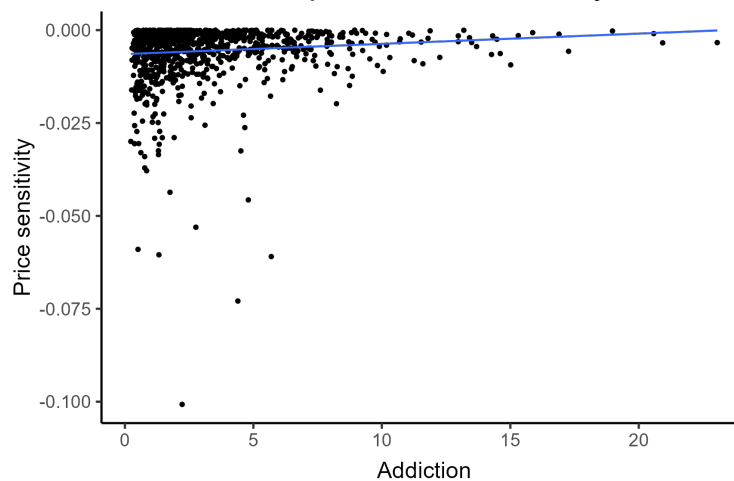
Beer and Others (-4.17 and -3.94) are higher than those for Beer and Happoshu (-1.71 and -2.09). These results potentially suggest a stronger impact of a liquor tax change on Third beer than that on Beer. Second, the cross-price elasticities between Third Beer and Other (0.512 and 0.465) are higher, indicating that these two categories are closer substitutes than other categories. Finally, the cross-price elasticity for Happoshu with respect to the Third beer (0.405) is also high, but the opposite direction is not (0.144). These observations

on substitution patterns are important for understanding the influence of category-specific liquor tax changes on demand.

Next, the comparison of the pre-and post-liquor tax change reveals that the overall price elasticities decrease, and more so for Happoshu, Third Beer, and Others. The lower price elasticities are mainly due to the decrease in overall alcohol consumption: despite a decrease in the Beer tax, the tax change lowered the overall addiction level, which in turn reduced alcohol consumption. We also observe changes in cross-price elasticities. For example, the cross-price elasticity for Happoshu with respect to Third Beer increases from 0.144 to 0.493, suggesting that due to an increase in the tax for Third Beer, Third Beer became a close substitute for Happoshu as well. A similar pattern is observed for Beer and Third beer.

Lastly, Figure 10 displays a plot of our estimates on price sensitivity and addiction levels per household. Our estimates show that households with higher levels of addiction tend to be less sensitive to changes in price. The implications of this result are significant, as it suggests that taxation aimed at reducing addiction may not be effective enough in motivating less-price-sensitive households. Therefore, policymakers might need to take into account the impact of addiction levels when designing pricing strategies and targeting specific consumer groups, which will be discussed in the counterfactuals.

Figure 10 Estimates: Relationship between Price Sensitivity and Addiction level



Notes: Each dot indicates a household in our estimation samples.

8. Counterfactual Simulations

8.1. Evaluating Planned Policy Change

Using the structural parameter estimates, we evaluate the impact of the liquor tax change. In the first counterfactual experiment, we evaluate the total impact by considering a scenario with no tax change at all and compare it to the baseline situation that follows the actual implementation schedule. In the second experiment, we consider an alternative tax policy change. In the actual implementation, tax levels are changed at multiple stages, starting in October 2020 until October 2026 (Figure 1). In the experiment, we consider a policy that changes all the tax rates to the final levels in October 2020. To simulate the quantities of our interest, we simulate consumer purchase decisions up to October 2027, one year after the last liquor tax change. Prices are simulated from the estimated price processes from the beginning of the sample period to Oct. 2027, and appropriate tax changes are reflected in the utility function (Equation 9). We thus assume a 100% pass-through of the tax changes to prices. We simulate D draws of price sequences, together with simulated store visits and idiosyncratic errors for choice, and fix the draws across all simulations.

The top panel of Table 14 shows the results for the first experiment with no tax change. As we expected, removing the tax change decreases Beer sales (-7.56%) and increases Third beer sales (16.56%). Also, we observe that sales for Happoshu and Other increase slightly (0.70% and 0.67%, respectively). Similar patterns are observed for firm revenue. Overall, the total change in sales across all categories is 1.44%, i.e., the tax change decreases the total sales slightly. The total change in firm revenue is smaller (0.18%), and the magnitude in Japanese Yen is negligible (JPY 0.08 per consumer and week). The tax revenue decreases due to the removal of the tax policy change (-0.69% or JPY -0.09). Due to the increase in alcohol purchases (and consumption), the average addiction level across consumers and weeks also increases by 1.72%. In other words, the tax policy change was effective at reducing consumers' alcohol addiction.

The bottom panel of Table 14 shows the results for the second experiment where the government changes all the taxes in October 2020, instead of changing them at multiple stages. Compared to the baseline scenario, the sales for Beer increases (9.37%), and the sales for the other three categories decreases (-6.65%, -6.10%, and -6.63%, respectively). This is because in the counterfactual scenario, Beer enjoys a lower tax rate immediately in October 2020 (in the actual schedule, it was lowered gradually from JPY 77 to JPY

Table 14 Counterfactual Experiment Results

	Beer	Happoshu	Third-beer	RTD	All
<i>CF1: No liquor tax change</i>					
Δ in quantity sold	-7.56%	0.70%	16.56%	0.67%	1.44%
Δ in firm revenue (%)	-7.40%	0.64%	14.25%	0.51%	0.18%
Δ in firm revenue (JPY)	-1.34	0.04	1.33	0.06	0.08
Δ in tax revenue (%)	2.92%	-0.45%	-10.47%	-1.21%	-0.69%
Δ in tax revenue (JPY)	0.19	-0.01	-0.24	-0.03	-0.09
Δ in addiction level (%)					1.72%
<i>CF2: Tax change all at once</i>					
Δ in quantity sold	9.37%	-6.65%	-6.10%	-6.63%	-1.41%
Δ in firm revenue (%)	9.11%	-6.30%	-5.66%	-5.55%	0.27%
Δ in firm revenue (JPY)	1.66	-0.34	-0.53	-0.66	0.12
Δ in tax revenue (%)	-1.88%	0.18%	6.65%	4.55%	1.17%
Δ in tax revenue (JPY)	-0.12	0.003	0.15	0.12	0.16
Δ in addiction level (%)					-1.61%

Notes: JPY changes in firm revenue and tax revenue are measured at consumer-week level.

54.25), but other categories suffer from a higher tax rate (JPY 54.25). Overall, this new tax policy change decreases the total sales by 1.41%. The change in firm revenue at the category level shows similar increasing and decreasing patterns, but the total change is slightly positive (0.27%), mainly due to a higher price for Beer than for the other three categories. The overall tax revenue also increases by 1.17%. Interestingly, compared to the baseline scenario, the addiction level decreases by 1.61%. As we saw, the immediate tax change for all categories decreases the category sales, which lowers the consumption of alcohol level and addiction levels. This is mainly due to the fact that consumers do not stockpile a lot before each policy change when the reform happens at once. Overall, this experiment suggests that if the objective of tax reform is to reduce the addiction to alcohol products and raise tax revenue, this alternative policy seems more ideal.

9. Conclusion

This paper studies the effects of alcohol taxes and regulations on alcohol consumption and addiction. In particular, we consider how the effectiveness of taxes and regulations for curbing addictive alcohol consumption vary by the degree of consumer present-bias. We study the research question with the consumer-level alcohol purchase data from Japan, where we observe multiple policy changes. By estimating a dynamic structural model of alcohol purchase and consumption with present-bias, we find that consumers are addictive to alcohol, and more so when they consume beverages with higher alcohol content.

We also find that consumers are forward-looking, while they are also present biased. A series of counterfactual simulations show that the current alcohol tax system in Japan makes consumers more addictive than the proposed change that equalizes tax rates across categories.

References

- Abbring J, Daljord Iskhakov F (2018) Identifying present-biased discount functions in dynamic discrete choice models . 40
- Abbring JH, Daljord Ø (2020) Identifying the discount factor in dynamic discrete choice models. *Quantitative Economics* . 40, 41
- Aguilar A, Gutierrez E, Seira E (2021) The effectiveness of sin food taxes: Evidence from Mexico. *Journal of Health Economics* 77:102455. 6
- Ailawadi KL, Neslin SA (1998) The effect of promotion on consumption: Buying more and consuming it faster. *Journal of Marketing Research* 35(3):390–398. 25
- Auld MC, Grootendorst P (2004) An empirical analysis of milk addiction. *Journal of Health Economics* 23(6):1117–1133. 18
- Bai L, Handel B, Miguel E, Rao G (2021) Self-control and demand for preventive health: Evidence from hypertension in India. *Review of Economic Statistics* 103(5):835–856. 40
- Becker GS, Grossman M, Murphy KM (1994) An empirical analysis of cigarette addiction. *The American Economic Review* 84(3):396. 3, 6, 18
- Becker GS, Murphy KM (1988) A theory of rational addiction. *Journal of Political Economy* 96(4):675–700. 6, 18
- Bollinger B, Sexton S (2018) Local excise taxes, sticky prices, and spillovers: Evidence from Berkeley's soda tax. *mimeo.* . 6
- Chan MK (2017) Welfare dependence and Self-Control: An empirical analysis. *Review of Economic Studies* 84(4):1379–1423. 4, 40, 41, 42, 43
- Chen J, Rao VR (2020) A dynamic model of rational addiction with stockpiling and learning: An empirical examination of e-cigarettes. *Management Science* 66(12):5886–5905. 6
- Ching AT, Osborne M (2020) Identification and estimation of Forward-Looking behavior: The case of consumer stockpiling. *Marketing Science* 39(4):707–726. 40
- Conlon CT, Rao N (2015) The price of liquor is too damn high: Alcohol taxation and market structure. 7
- Conlon CT, Rao NL (2020) Discrete prices and the incidence and efficiency of excise taxes. *American Economic Journal: Economic Policy* 12(4):111–143. 7

- Dragone D, Raggi D (2021) Resolving the milk addiction paradox. *Journal of Health Economics* 77:102452. [6](#), [18](#), [19](#)
- Dubois P, Griffith R, O'Connell M (2020) How well targeted are soda taxes? *American Economic Review* 110(11):3661–3704. [6](#)
- Erdem T, Imai S, Keane MP (2003) Brand and quantity choice dynamics under price uncertainty. *Quantitative Marketing and Economics* 1(1):5–64. [31](#), [32](#), [33](#)
- Fang H, Silverman D (2009) Time-inconsistency and welfare program participation: Evidence from the nlsy*. *International Economic Review* 50(4):1043–1077. [34](#), [40](#), [42](#)
- Fang H, Wang Y (2015) Estimating dynamic discrete choice models with hyperbolic discounting, with an application to mammography decisions. *International Economic Review* 56(2):565–596. [4](#), [35](#), [43](#)
- Ferrall C (2012) Explaining and forecasting results of the self-sufficiency project. *Review of Economic Studies* 79(4):1495–1526. [43](#)
- GBD 2016 Alcohol Collaborators (2018) Alcohol use and burden for 195 countries and territories, 1990–2016: a systematic analysis for the global burden of disease study 2016. *The Lancet* 392(10152):1015–1035. [55](#)
- Gordon BR, Sun B (2015) A dynamic model of rational addiction: Evaluating cigarette taxes. *Marketing Science* 34(3):452–470. [4](#), [6](#), [16](#), [17](#), [22](#), [29](#), [33](#)
- Griffith R, O'Connell M, Smith K (2019) Tax design in the alcohol market. *Journal of Public Economics* 172:20–35. [6](#)
- Grossman M, Chaloupka FJ, Sirtalan I (1998) An empirical analysis of alcohol addiction: Results from the monitoring the future panels. *Economic Inquiry* 36(1):39–48. [6](#)
- Gruber J, Köszegi B (2001) Is addiction “rational”? theory and evidence. *The Quarterly Journal of Economics* 116(4):1261–1303. [3](#), [4](#), [20](#), [41](#)
- Hendel I, Nevo A (2006a) Measuring the implications of sales and consumer inventory behavior. *Econometrica* 74(6):1637–1673. [33](#), [40](#)
- Hendel I, Nevo A (2006b) Measuring the implications of sales and consumer inventory behavior. *Econometrica* 74(6):1637–1673. [33](#)
- Hendel I, Nevo A (2006c) Sales and consumer inventory. *The RAND Journal of Economics* 37(3):543–561. [3](#), [22](#)
- Hinnosaar M (2016) Time inconsistency and alcohol sales restrictions. *European Economic Review* 87:108–131. [3](#), [7](#), [25](#), [26](#)
- Hitsch GJ, Hortaçsu A, Lin X (2021) Prices and promotions in U.S. retail markets. *Quantitative marketing and economics* 19(3-4):289–368. [23](#)

- Hollenbeck B, Uetake K (2021) Taxation and market power in the legal marijuana industry. *The RAND Journal of Economics* 52(3):559–595. 7
- Imai S, Jain N, Ching A (2009) Bayesian estimation of dynamic discrete choice models. *Econometrica* 77(6):1865–1899. 4, 36, 37
- Ishihara M, Kwon M, Mizuno M (2022) An empirical study of scarcity marketing strategies: Limited-time products with umbrella branding in the beer market. *Forthcoming in Journal of the Academy of Marketing Science* . 43
- Khan R, Misra K, Singh V (2016) Will a fat tax work? *Marketing Science* 35(1):10–26. 6
- Kim JY, Ishihara M (2021) Evaluating the effect of soda taxes using a dynamic model of rational addiction. *Available at SSRN 3892506* . 4, 6, 16, 22, 29
- Kong X, Dubé JP, Daljord 2022) Non-parametric estimation of habitual brand loyalty URL <http://dx.doi.org/10.2139/ssrn.4070747>. 41
- Kwon M, Erdem T, Ishihara M (2021) Counter-cyclical price promotion: Capturing seasonal category expansion under endogenous consumption. *Working Paper* . 33
- Laibson D (1997) Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics* 112(2):443–477. 40
- Levine J (2022) Are menthol cigarettes more addictive? a Cross-Category comparison of habit formation. 58
- Magnac T, Thesmar D (2002) Identifying dynamic discrete decision processes. *Econometrica* 70(2):801–816. 40
- Miller KM, Sahni NS, Strulov-Shlain A (2022) Sophisticated consumers with inertia: Long-term implications from a large-scale field experiment. *Available at SSRN 4065098* . 40
- Miller NH, Weinberg MC (2017) Understanding the price effects of the millercoors joint venture. *Econometrica* 85(6):1763–1791. 7
- Miravete EJ, Seim K, Thurk J (2023) Welfare consequences of nominal excise taxation? 54
- O’Donoghue T, Rabin M (2006) Optimal sin taxes. *Journal of Public Economics* 90(10-11):1825–1849. 4, 34
- Piccoli L, Tiezzi S (2021) Rational addiction and time-consistency: An empirical test. *Journal of Health Economics* 102546. 6
- Schmacker R, Smed S (2023) Sin taxes and Self-Control. *American Economic Journal: Economic Policy* 15(3):1–34. 6
- Seiler S, Tuchman A, Yao S (2021) The impact of soda taxes: Pass-Through, tax avoidance, and nutritional effects. *Journal of Marketing Research* 58(1):22–49. 6
- Tsubota T (2021) Identifying dynamic discrete choice models with hyperbolic discounting . 40
- Wang C, Weiergraeber S, Xiao R (2022) Identification of dynamic discrete choice models with hyperbolic discounting using a terminating action. 40

Wang EY (2015) The impact of soda taxes on consumer welfare: implications of storability and taste heterogeneity. *The RAND Journal of Economics* 46(2):409–441. [6](#), [33](#)

Appendix

A. Data Construction

A.1. SKU to Brand Definition

The raw data defines the product at the SKU level, and the number of total unique SKUs in our raw data is 21,564. To analyze, we first aggregated SKUs into brands. The details of this procedure are as follows.

First, we define the *brand* of each SKU relying on its description documented in the original purchase data. The description contains 1) the brand name, 2) auxiliary information (flavor, ingredient, campaign, etc.), 3) package type (can, bottle, etc.), 4) the unit size of a package (ML), and 5) the number of units (6 pack, 24 pack, etc). The distinction between 1) and 2) are often ambiguous. Even if two SKUs' descriptions partially share the same brand name, we separate those into different brands based on price tiers (e.g., Super-Dry and Super-Dry Premium), main product features (Tanrei Green Label and Tanrei Platinum Double), and marketing strategies (Ichiban-Shibori and Ichiban-Shibori local limitations), etc. In contrast, we combine the SKUs that share the same brand name with different flavors as the same brand to avoid too much proliferation of brand names, particularly for non-beer, easy-to-drink alcoholic beverages.

Second, we create "meta brands" by integrating some brands in order to match the advertising exposure data. Although both data were collected from the same individuals, their brand definitions are inconsistent. In the advertising data, brands are often separated by advertising campaigns. We reduced the variation of brands to make the brand definitions consistent for both data.

A.2. Product Characteristics

To obtain the product characteristics for each brand, we first scraped every alcohol product's detailed description from each manufacturer's website. All the major beer brands show up on the official sites. The information contains the degree of alcohol, calorie, protein, fat, sugariness, carbon, fiber, salt, and prin. In several cases, we did not observe values in some specific characteristics such as salt while we always observe the degree of alcohol.

There are two difficulties in merging this directly with the purchase dataset. First, we observe the JAN code only for each product for Kirin and Suntory. We did not observe the JAN code for Asahi and Sapporo. Second, we cannot observe all the old products. Data are collected in August 2021, and information showing up on the homepage is limited to only the currently sold products. For products sold previously but stopped production prior to this scraping timing, we could not obtain product characteristics from the official homepage. These prevent us from obtaining product characteristics for every SKU level. This is a demanding task as SKU varies just by flavors and/or tiny model changes which do not affect most of the nutritional characteristics. Usually, it is challenging to collect all information for past products at the SKU level. Therefore, we did not merge by SKU level but by brand level. For our definition and the construction of the brand, see section xxx.

At the brand level, we confirmed that the main characteristics, especially the alcohol content, rarely change over time within the same brand according to the press release of each manufacturer (later, we also verified this using reliable third-party information). This allows us to use the same product characteristics for every

SKU in the same brand for any period. This dramatically reduces the missing information for any SKUs as long as one of the SKUs in the same brand has the official information of the product characteristics and that SKU is the representative of the brand. In most cases, this condition is satisfied well and we can fill in the missing characteristics.

Still, we had some exceptions for some brands. First, for a brand in which we could not observe product characteristics for any SKUs from the manufacturer's official website, we searched package pictures using google search and [amazon.co.jp](https://www.amazon.co.jp). We could find the picture with the nutrients side and use these values. Second, there are several brands that have different product characteristics within the SKUs. This never is only case for a few ready-to-drink brands, and not the issue for beer brands. This is just because there are more flavors and minor regional differences of the promotion in ready-to-drink than beer products. In such a case, we used the product characteristics of the most sold SKU within the brand as a representative SKU of the brand.

Lastly, we verified the credibility of these values, especially for the alcohol content and its time variation. To do so, we used a few third-party databases such as <https://www.eatsmart.jp/>, <https://shareview.jp/>, and google image searches. This is useful to verify our recent characteristics match with the older information as the updated date tends to be older in those sites. We never observe the variation in the alcohol content. For some of the other characteristics such as salt, we sometimes observe the unique values in these third-party databases while we saw the interval values (such as 0.1-0.3mg) in the official information. In such a case, we use the median value (0.2mg). Also, some information from the official site contains the approximated values (such as approx. 2.0mg), then we just use the value (2.0mg) directly. This completes the product characteristics' construction.

A.3. Sample Selection

In analysis, we limit the focus on 350ML and 500ML and 1pack, 6pack, and 24pack as they are the majority of our dataset. The top panel of the table 15 shows that 350ML and 500ML are major sizes and share almost 99% of the purchase. Also, by investigating the purchasing pattern by households, 38.2% of households purchased only 350ML size. 58.8% households purchased both 350ML and 500ML at least once during the sample period.

The bottom panel of the table 15 shows that the major pack is 1 pack, 6 pack, and 24 packs in order, which accounts for 99.4% share of the total purchase. Also, by investigating the purchasing pattern by households, 33.3% of households purchased only single packs, 4.1% of households only purchased 6 packs, and 1.0% of households purchased only 24 packs. 25.1% of households purchased both 1 pack, 6 pack, and 24 pack at least once during the sample period.

B. Additional Descriptive Evidence

B.1. Tax Pass-Through Rate

Move the main to here if necessary. If not, delete this section.

B.2. Effect of Lockdown due to COVID19 on Alcohol Products' Purchase

TBA

Table 15 Summary Statistics of Size and Packs

UnitSize	135	240	250	275	303	305	310	320	330	334	340	350	490	500	633					
Freq	0.003	0	0.003	0	0	0.002	0	0	0.002	0	0	0.724	0	0.263	0.003					
Pack	1	1.42	2	3	4	4.28	5	6	7	8	8.57	9	10	12	14	17.1	18	20	24	48
Freq	0.714	0.004	0	0	0	0	0	0.225	0	0	0.001	0	0	0	0	0	0	0	0.055	0

Table 16 Effect of COVID19 Lockdown on Alcohol Product Purchase

	(1)	(2)	(3)
lockdown _t	0.015 (0.004)	-0.002 (0.004)	-0.007 (0.009)
log(<i>Price</i>) _{it}		-0.712 (0.037)	-0.888 (0.289)
IV			✓
Year FE	✓	✓	✓
Month FE	✓	✓	✓
Household FE	✓	✓	✓
R ²	0.51	0.53	0.53
Observations	389,758	389,758	389,758

Note: The dependent variable is $\log(Q_{it})$. The unit of observation used in regression is household-date-level. Year-Monthly FE and household FE are included.

B.3. Effect of Tax Policy Change on Advertising Exposure Decisions

To investigate whether manufactures increased the advertising prior to the alcohol tax policy change at 2020/10. If so, consumption can increase as a response to more advertising.

TBA

TBA

$$\text{Ads exposure}_{ct} = \mu_c + \mu_t + \epsilon_{it}.$$

TBA

C. Details of Structural Estimation

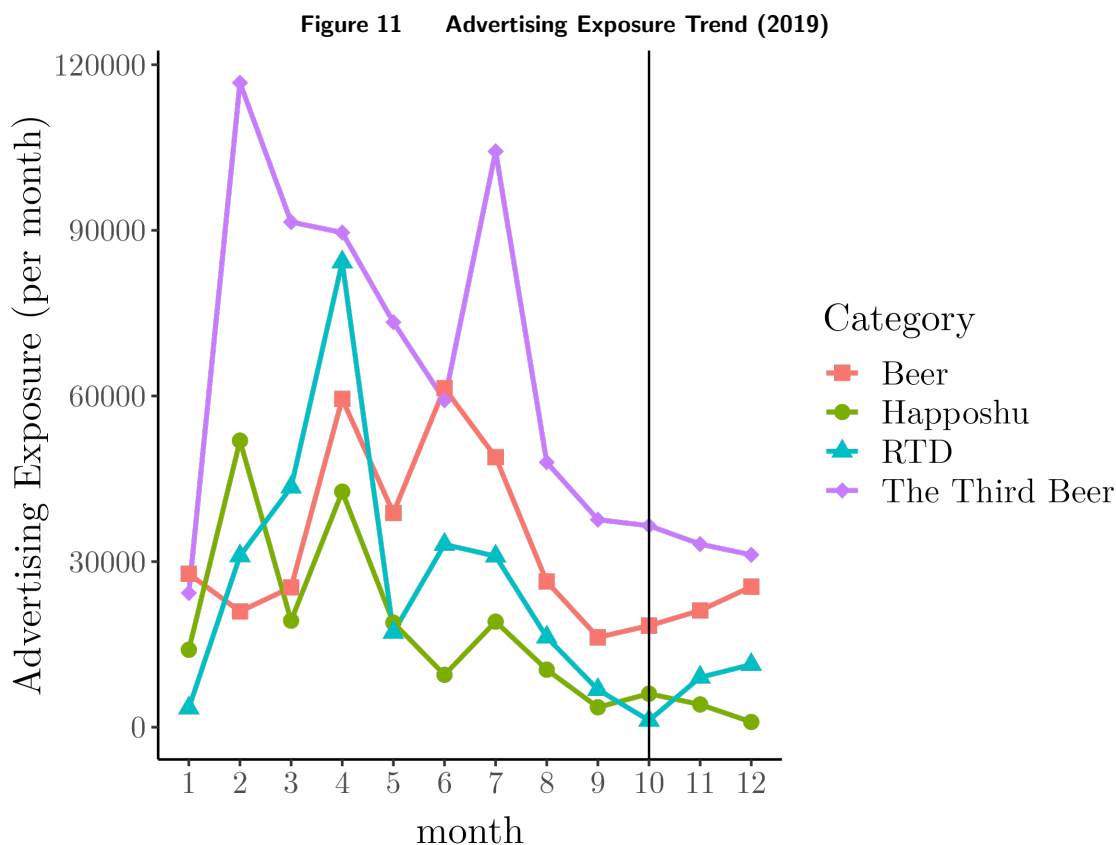
TBA

C.1. Instruments for Inventory

TBA

D. Mapping Alcohol Consumption into Mortality

This section details how we convert the consumption change into the mortality changes. Our procedure follows [Miravete et al. \(2023\)](#) and uses the estimates on alcohol consumption and the relative risk of disease



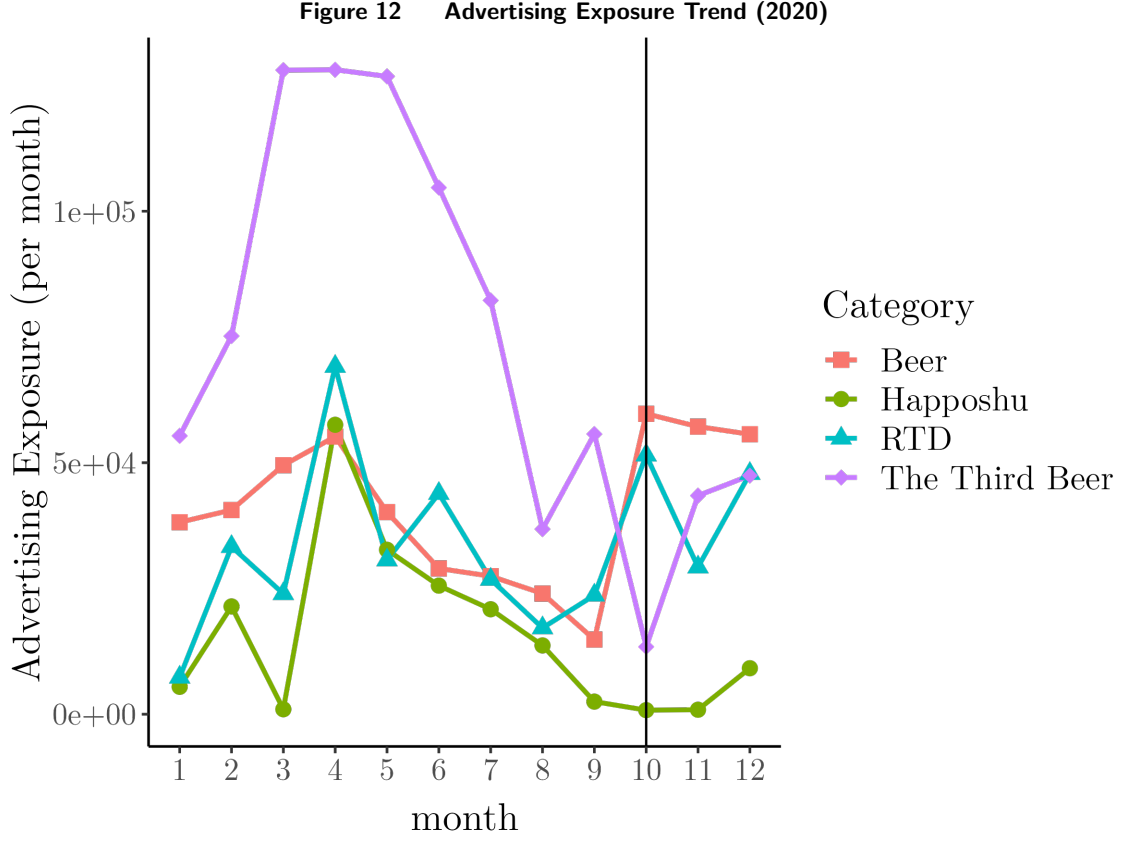
Notes:

Table 17 Instruments Relevance

	$\log(\text{Purchase Quantity}_{st})$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	11.8 (3.18)							
$\log(\text{elec})$	-0.820 (0.355)	-0.854 (0.383)	-2.09 (0.878)	-2.09 (0.992)	-0.152 (0.111)	-0.170 (0.043)	-0.163 (0.051)	-0.540 (0.126)
R^2	0.02	0.04	0.05	0.07	0.83	0.78	0.82	0.79
Observations	5,148	5,148	5,148	5,148	5,148	5,148	5,148	5,148
Year fixed effects		✓		✓	✓		✓	
Month fixed effects			✓	✓	✓			✓
Prefecture fixed effects					✓	✓	✓	✓

Note: The unit of observation used in the regression is weekly/prefecture-level. The DV is Alcohol purchase quantity_{st}. We included household FE, and year/month FE in all the regressions. SEs are in parenthesis and clustered by prefecture.

from the recent epidemiology study by [GBD 2016 Alcohol Collaborators \(2018\)](#). The study reviews XXX and provides data on the relative risk for multiple diseases such as lung cancer and heart disease.



Notes:

E. Identifying Moments

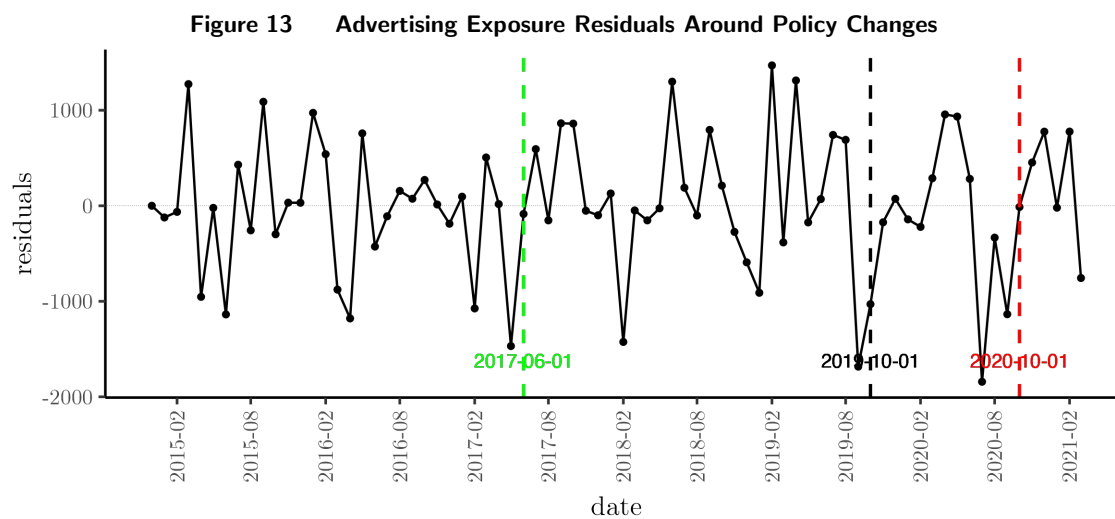
$$\log\left(\frac{p_{k,t}(x)}{p_{K,t}(x)}\right) = u_{k,t}(x) + \beta\delta [Q_k(x) - Q_K(x)] \left[m_{t+1} + \sum_{\tau=t+2}^T \delta^{\tau-t-1} \left(\prod_{r=t+1}^{\tau-1} Q_r^{pb}(\beta) \right) m_\tau \right] \quad (20)$$

$$\begin{aligned} \log\left(\frac{p_{k,t}(x)}{p_{K,t}(x)}\right) - \log\left(\frac{p_{k,t}(x')}{p_{K,t}(x')}\right) &= \beta\delta [Q_k(x) - Q_k(x') - Q_K(x) + Q_K(x')] \\ &\times \left[m_{t+1} + \sum_{\tau=t+2}^T \delta^{\tau-t-1} \left(\prod_{r=t+1}^{\tau-1} Q_r^{pb}(\beta) \right) m_\tau \right] \end{aligned}$$

- Drive $z_{it} \rightarrow \infty$, then $H(z_{it}) \rightarrow \infty$, which implies the consumption is equal to the purchase amount.
- Note that the addiction (remaining latent variable) is a deterministic function of the consumption path.

Hence, conditioning on the consumption path is equivalent to conditioning on the addiction level.

- Conditioning on $C_{t=1}^{\tau-1}$ is problematic because there is no variation in time till a policy change. We need to compare two consumers with the same consumption path (and hence the same addiction level), but the time till a policy change (excluded variable) is different.
-



Notes: The figure shows how the residual parts of the alcohol consumption regression change over time. The alcohol consumption regression uses household-month level data on alcohol consumption amount (mg), and we regress alcohol consumption amount on price index, the household fixed effects, and year-month fixed effects. The price index is constructed by taking the weighted mean of the price in scanner data for each prefecture, year, and month. We instrumented the Price index by the VAT rate, dummies for discount regulation in 2017/06, and alcohol tax policy changes in 2020/10. The green vertical dotted line indicates the discount regulation in 2017/06, the black vertical dotted line indicates the VAT reform in 2019/10, and the red vertical dotted line indicates the alcohol tax policy reform in 2020/10.

Online Appendix

OA.1. Conceptual Model of Addiction

- Before developing a full dynamic structural model, we show a conceptual framework and evidence of addictive consumption and discuss why we need to incorporate the addictive aspects in our model.
- [Levine \(2022\)](#) shows a simple conceptual framework that provides a testable prediction against the null hypothesis that there is no state dependence.
- Among the state dependence, we refer it as the addiction. Note that the state dependence is consistent with both the addiction and habitual formation models.
- Adapting the idea, we demonstrate how past consumption levels could affect consumers' alcohol product choices, and why incorporating this feature is important for policy-making.
- For brevity, we abstract away the dynamic aspect of consumption choice other than addiction although we will incorporate it in our full model.

We assume that the current category consumption follows the following data-generating process:

$$c_t = (1 - \lambda)\mu + \lambda a_t - \epsilon_t, \quad (21)$$

where μ is a consumer's static preferences for an alcohol product and a_t is a current level of addiction. $\lambda \in [0, 1]$ captures the weight of the fully static model vs the addiction model. These represent that if consumers are addicted ($\lambda > 0$), they would consume more than they would do without addiction. ϵ_t is an iid random shock with mean zero. Note that we omit the consumer's subscript i for this model as everything is at the consumer level.

The addiction level changes based on the consumer's current level of consumption, and we specify as follows:

$$a_{t+1} = (1 - \kappa)a_t + \kappa c_t, \quad (22)$$

where κ is a depreciation rate of the addiction level, capturing how the current consumption level affects the consumer's next period's addiction level.

At the static baseline where no addiction exists, the steady-state consumption level is $\bar{c} = (1 - \lambda)\mu$. On the contrary, the steady-state consumption level is $\bar{c} = \bar{k} = \mu$ if addiction exists. Thus, the difference in consumption level between the two cases is summarized in $(1 - \lambda)$, and the difference in consumption level is larger with the higher degree of addiction ([Levine 2022](#)). This triggers an issue in identifying addiction parameters (λ, κ) because you cannot distinguish them at the steady state consumption level μ . In other words, a different set of values on (λ, κ) would predict the same value of consumption at the steady state of addiction, failing to find a model.

Instead, we focus on the exogenous shocks that shift the consumption level out of the steady-state level following [Levine \(2022\)](#). Specifically, we use the tax increase as a shock that shifts the consumption level higher than the steady-state level. To demonstrate how the identification strategy works, suppose there are two identical consumers at steady state, who are randomly assigned to keep or increase consumption level at a period prior to tax increase.

We denote the consumption level at period t as $c_t(x)$, where t represents the period after the exogenous consumption hike in one of the consumers and x indicates the treatment variable which is equal to 1 for a consumer with consumption hike and 0 for another consumer. Therefore, the consumption level at period 1 for each consumer can be written as follows:

$$\begin{aligned} c_1(1) &= (1 - \lambda)\mu + \lambda a_1(1) \\ c_1(0) &= (1 - \lambda)\mu + \lambda a_1(0) \end{aligned} \tag{23}$$

By subtracting the two equations, we obtain the following:

$$\begin{aligned} \Delta c_t &:= c_t(1) - c_t(0) \\ &= \lambda \Delta a_t \end{aligned} \tag{24}$$

This equation means that there is no difference in consumption level if $\lambda = 0$ while there is a consumption level difference between a consumer who is exposed to an exogenous consumption hike and a consumer who did not if addiction exists.

$$\begin{aligned} \Delta a_1 &:= a_1(1) - a_1(0) \\ &= (1 - \kappa)(a_0(1) - a_0(0)) + \kappa(c_0(1) - c_0(0)). \end{aligned} \tag{25}$$

Note that $\Delta a_1 = 0$ under the assumption of the random assignment: $a_0(1) = a_0(0)$ and $c_0(1) = c_0(0)$. In this case, we cannot identify λ because $\Delta c_t = \lambda \cdot 0 = 0$. However, if we allow the non-random allocation (self-selection into the treatment) and instead assume the monotonicity ($a_0(1) > a_0(0)$ and $c_0(1) > c_0(0)$), we can identify λ because $\Delta a_1 > 0$ so $\Delta c_t = \lambda \Delta a_t \geq 0$. Importantly, these assumptions can be testable in the observed data. To test whether addiction exists in the sample, all we need is to test against $\Delta c_t = 0$.

Although this conceptual framework can be easily extended to the case with multiple categories and products, we still oversimplify the following important aspects of consumers' alcohol product choices.

- Consumers do not stockpile.
- Consumers have unlimited holding space without any cost.
- Consumers are myopic and do not rationally take the future value (saving, holding cost, addiction) into account.