RIETI Discussion Paper Series 20-E-092

Three Minds Equal Manjushari's Wisdom: An Anatomy of Informal Social Learning with Heterogenous Agents by the Hierarchical Bayesian Approach

SATO, Masahiro
Tohoku University

OTA, Rui
Yokohama City University

ITO, Arata
RIETI

YANO, Makoto
RIETI

RIETI
Research Institute of Economy, Trade & Industry, IAA

The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 20-E-092
December 2020

Three Minds Equal Manjushari's Wisdom: An Anatomy of Informal Social Learning with Heterogenous Agents by the Hierarchical Bayesian Approach[1]

SATO, Masahiro (Tohoku University)
OTA, Rui (Yokohama City University)
ITO, Arata (Research Institute of Economy, Trade and Industry)
YANO, Makoto (Research Institute of Economy, Trade and Industry)

Abstract

We learn from all sorts of informal social learning devices, which convey information only inaccurately. Despite this, however, a case supporting a positive contribution of such a device has not been captured in the existing empirical literature. This study builds a discrete choice model of consumption in which informal social learning takes place in a Beta-Bernoulli process of information update. The model is estimated by the Bayesian statistical method with Markov chain Monte Carlo simulation. It provides evidence supporting the positive role of an informal device, to which individual heterogeneity and the effacing of bad news contribute.

Keywords: social learning, consumer, Bayesian learning, hierarchical Bayesian estimation, discrete choice model
JEL classification: C11, C35, D12, D83, O33

[1] This study is conducted as a part of the Project "Evidence-based policy study on the law and economics of market quality" undertaken at the Research Institute of Economy, Trade and Industry (RIETI). This study is supported by JSPS KAKENHI Grant Numbers JP16H02015 and JP19K13674. The authors are grateful for helpful comments suggestions by Discussion Paper seminar participants at RIETI.

1. Introduction

Social learning is the act of individuals to learn from one another on a similar level. There are many devices that specifically aim to facilitate social learning. Conventional examples are local pages in newspapers and local weeklies. More modern examples are reviews that are posted at online shopping and/or specialized internet sites. While these devices offer micro signals from which one can obtain other individuals' respective views, TV ratings and average stock price indices are macro signals that convey information on the perceptions and actions of a society as a whole.

There are many such formal devices for social learning, which are to dissimilate, or at least are designed to dissimilate, accurately information to the society. However, we learn a lot more from all sorts of informal devices, including the way in which others use new gadgets, the way in which others dress, and the way in which others react to certain materials and ideas, and so on. Despite the importance, however, the effectiveness of informal social learning has not been evidenced in the existing empirical literature.

The main purpose of this study is to investigate if an informal device for social learning in fact plays a positive role in dissimilating private information throughout a society. With regard to social learning, there are at least two fundamental questions. The first is: Does social learning play a productive role guiding the society towards the truth? Second, if so, is social learning led by a particular type of people or merely by random factors; or, more specifically, does the heterogeneity of individuals contribute to social learning?

We address these questions with respect to informal social learning devices. In doing so, we focus on a frozen food market that suffered from a serious but idiosyncratic product defect; an idiosyncratic event can be defined as an event that is non-essential and could, and should, be ignored (see Honryo and Yano, 2020). Even if the product defect is purely idiosyncratic, at first, consumers might believe otherwise, in which case they would, suboptimally, stay away from the product until they learn that the defect is in fact idiosyncratic. For such an event, store shelves may be thought of as an informal social learning device, providing consumers with some information, if inaccurate, on whether or not other consumers suffer from a defect.

In order to capture the role of informal social learning, we develop a new empirical model with fully heterogenous individuals by incorporating a Bayesian learning process into a standard discrete choice model (McFadden, 1973, and Train, 2003). The structural estimation of our model provides some evidence supporting that informal social learning (from products lineups on store shelves) plays a positive role to the recovery from a negative shock by selling a product with an idiosyncratic defect; the heterogeneity of individuals appears to contribute to this process of recovery.

Into this model, we incorporate an aspect of social learning in a broader sense that may take place through a formal device, i.e., newspaper reports related to the product defect.
[2] Our estimation results suggest that more precisely, the more readily people efface the bad news reported by the formal device, the more likely they resume purchasing products of the company that suffered from an idiosyncratic negative shock.

This study demonstrates that the Beta-Bernoulli process is useful for capturing social learning through the interaction of heterogenous individuals in an empirically tractable manner. In that process, the Bayesian update is characterized simply by two simple macro signals: The total number of trials (or that of people who buy products of the company that sold a defective product) and that of realizations of a particular event (the number of people who actually buy, and suffer from, a defective product of that company). If those numbers were publically available, formal social learning would take place. In a consumer market, however, the number of people who buy products of a particular company is usually not available; even if such a number were announced by the company that sold a defective product, it would be unlikely that consumers would believe the number. This study provides evidence supporting that even in such a situation, store shelves serve as a learning device, facilitating informal social learning. While the use of a Beta-Bernoulli distribution has not been recognized in the empirical literature on social learning, it is common in the fields of machine learning (e.g. Chen et al., 2013; Lee and Hong, 2016; Akhtar and Mian, 2018). Israel (2005) and Tomlin (2009) adopt it for the analysis of personal learning (learning from one' own experience in the analyses of consumer or corporate behaviors; Shen and Djurić (2014) and Zhao and Sayed (2015) develops its use in simulation, and Davis, Gaur and Kim (2020) in experimental analysis.

In order to capture the roles of an informal social learning device, we adopt the hierarchical Bayesian estimation with Markov chain Monte Carlo (MCMC) simulations; we characterize an individual preference by 8 different parameters. In our estimation, we use daily scanner panel data representing barcode-level purchase records over some 680,000 items for about 50,000 survey participants in Japan; this data covers the defective frozen food incidence that occurred in 2013. Although, in order to render estimation tractable, we focus on 100 individuals who purchased related frozen food items most frequently over the period, we have at least $800 = 8 \times 100$ parameters to estimate. This shows the use of the MCMC in estimating a large number of parameter values from data. The hierarchical Bayesian method makes it possible to conduct a structural estimation not only for the parameters that characterize the

[2] In the existing literature, social learning is thought of in a narrower sense, referring to learning facilitated by observations of, or interactions with, another individual (or its products) on a similar level; see Hoppitt and Laland (2013).

preferences of all respective individuals but also the underlying hyperparameters. This method has been widely adopted in the recent literature on personal learning or the process in which an individual learns from his own personal experiences; see Akçura, Gönül, and Petova (2004), Iyengar, Ansuri, and Gupta (2007), Narayanan and Manchanda (2009), and Zhao, Zhao, and Helsen (2011).[3]

In the theoretical literature, Banerjee (1992) and Bikhchandani and Hirshleifer and Welch (1992) model a "formal learning device" by that agents can observe a full history of their predecessors' actions. Vives (1993) studies an "informal social learning device" by which the aggregate action of others is observable through noisy prices. Smith and Sørensen (2000), Chamley (2003), Banerjee and Fudenberg (2004), Herrera and Hörner (2013), Besbes and Scarsini (2018) study "informal social learning devices" by which the actions of others are partially observable. The informal social learning device on which we focus in the empirical context (i.e., store shelves) is "far more informal" in that it provides much more noisy and partial observations on others' actions.

There are some empirical studies that are concerned with social learning with a formal device and/or identical agents. Roberts and Urban (1988) study a process in which consumers who are identical except for personal experiences learn about the average quality of products from word-of-mouth information that are common to all consumers. Zhao, Yang, Narayan, and Zhao (2013) study a process in which potentially heterogeneous consumers learn from their own experiences and the average online reviews that are common for everyone (or a formal social learning device).

In what follows, in Section 2, we introduce informal social learning into a discrete choice model of consumption. In Section 3, we explain data and how we represent the variables in a model by data. We explain the estimation method in Section 4 and evaluate estimated parameters of the model in Section 5. In Section 6, we explain the roles of the informal device and the heterogeneity of individuals in social learning.

2. Model of Informal Social Learning

In this section, we build a discrete choice model of a market in which a supplier suffers from an idiosyncratic product trouble and in which consumers will gradually learn that the trouble is idiosyncratic, or that the product is safe. Social learning is modelled by a Beta-Bernoulli distribution, into which we incorporate the availability on store shelves of products of the troubled supplier as a social learning device. The more the supplier's products are available in stores, the more likely consumers learn that the defect is

[3] For recent economic applications of the hierarchical Bayesian analysis, see Meager (2016, 2019).

idiosyncratic, or the product is safe.

2.1. Idiosyncratic Pesticide Contamination Case

In this study, we focus on the Japanese frozen food market, in which a worker of a popular supplier, Aqli Foods, injected a poison into Aqli products that he handled. This incidence turned out to be idiosyncratic in that once the worker was caught, no similar poisoning incidence followed. Despite this, the demand for Aqli products decreased significantly; and it took a long time for Aqli Foods to recover from the incidence even partially. In order to isolate the effect of informal social learning, this incident provides an "ideal" working ground for several reasons. First, because the Aqli product poisoning was an isolated incidence and because no one knew at first why some products were contaminated, we may assume that all customer good was reset at the time at which the idiosyncratic defect is discovered; at that time, all individuals restarted learning. Second, we may assume that consumers lost their trust in Aqli so that they would heavily discount any information that Aqli provides; therefore, it is likely that consumers relied informal devices to learn about the safety of Aqli products. Third, because the defective product caused a serious health problem, it took a long time for many consumers to return to Aqli products.

In December 2013, Aqli Foods Corporation, a medium-size domestic food company, detected in its returned products malathion, a highly concentrated organophosphorus pesticide[4]. On December 29, Maruha Nichiro Holdings, the parent company of Aqli Foods, announced this fact, a voluntary recall of the suspected products, and the temporally closing the factory in the Gunma prefecture that produced the suspected products. Eventually, 67 household items and 45 business items ranging from croquettes, gratins, dorias, lasagnas, pizzas, pancakes, through pies were recalled. The list of recalled products was announced repeatedly in national and local newspapers starting December 30.

On January 25, 2014, a contract employee of Aqli's Gunma factory was arrested for having intentionally mixed the toxic agrochemicals into the products; Maruha Nichiro announced that the presidents of Maruha Nichiro and Aqli Foods were to take responsibility by resigning at the end of March. Maruha Nichiro also formed an independent committee to investigate the causes and background of the incident in January and organized another committee in April to propose solutions to reconstruct its crisis-management strategies and improve its food defense processes. On May 29, it published measures to prevent a recurrence, which included more monitoring cameras in the Gunma factory. The Gunma factory restarted its production step by step, starting with three of the five product categories in August, 2014, and the other two in October.

---

4 See Associated Press, 2014, Kyodo, 2014, Tatusian, 2014, and Viet Nam News, 2014.

2.2. Bigdata on consumers' purchases

In our empirical study, we use part of *Intage SCI* (Nationwide Consumer Panel Survey), which is daily scanner panel data staring in April 2010 and representing barcode-level purchase records for about 50,000 respondents in Japan.[5] Respondents scan the barcodes (Japan Article Number, JAN[6]) of products using a portable barcode scanner or smartphone at the time of purchase and enter the purchase details (where and when they bought the item, how much it cost, how much they spent shopping in total) by the end of the day. It contains more than 680,000 barcodes sold at supermarkets, convenience stores, 99 or 100 yen stores, hardware/discount stores, pharmacies/ drugstores, liquor stores, and department stores in 46 of Japan's 47 prefectures (excluding Okinawa). The major product categories covered are staple foods, processed foods, alcohol and non-alcohol beverages, household goods, cosmetics, and drugs. The data covers products sold by Aqli Foods and its competitors.

2.3. Heterogenous Informal Social Learning in the Beta-Bernoulli Process

In what follows, we introduce an estimable model of informal social learning with heterogeneous consumers facing an idiosyncratic health hazardous defect. Assume that there is a fixed number of consumers. They are heterogenous in both their preferences and the ways in which they form their respective beliefs with respect to the harm from a health hazardous product. Each consumer engages in Bayesian learning.

Let $\lambda_{nt} \in \{0, 1\}$ denote a realization of the random variable indicating the occurrence of a harmful event to individual n who buys a product of the troubled company in period t. Suppose λ_{nt} follows a Bernoulli distribution, $\lambda_{nt} \sim Ber(\theta)$, where θ be the probability with which an individual suffers from a product defect when he buys a product of the troubled company, that is, $\text{Prob}(\lambda_{nt} = 1) = \theta$. No individual knows this probability; however, he has a prior belief on θ update it through observations of others' actions and their outcomes. Assume that such learning of an individual follows a Beta-Bernoulli process. In period t, individual n has a conjugate prior on θ obeying a beta distribution, $\theta \sim Beta(\varphi_{nt}, \psi_{nt})$, with individual parameters $\varphi_{nt} > 0$ and $\psi_{nt} > 0$. The mean and variance of individual n's prior belief on θ are given by

[5] The number of respondents contains in the dataset grows step by step; about 20,000 respondents in April 2010 - March 2011, 27,000 in April 2011- December 2011, and about 50,000 in each of the years since January 2012. See https://www.intage.co.jp/english/service/platform/sci/ for more details.

[6] The JAN code is the product code employed in Japan and is used for barcode representation in point of sale (POS) systems, ordering systems, and inventory control systems. It forms part of the GTIN (Global Trade Item Number) system, the globally standardized product identification system for trade in retail and other supply chains, which includes the European Article Number (EAN) used in Europe and the Universal Product Code (UPC) employed in the United States and Canada.

$$E_{nt}[\theta] = \frac{\varphi_{nt}}{\varphi_{nt} + \psi_{nt}} \tag{1}$$

and

$$Var_{nt}[\theta] = \frac{\varphi_{nt}\psi_{nt}}{(\varphi_{nt} + \psi_{nt})^2(\varphi_{nt} + \psi_{nt} + 1)}. \tag{2}$$

Let $\widetilde{\mathrm{N}}_t$ be the set of individuals who purchase a product of the troubled company in period t. Then, it is known that individual parameters φ_{nt} and ψ_{nt} satisfy

$$\varphi_{nt} = \varphi_{n,t-1} + \sum\nolimits_{m \in \widetilde{\mathrm{N}}_t} \lambda_{mt} \tag{3}$$

and

$$\psi_{nt} = \psi_{n,t-1} + \widetilde{N}_t - \sum\nolimits_{m \in \widetilde{\mathrm{N}}_t} \lambda_{mt} \tag{4}$$

where $\widetilde{N}_t = \#\widetilde{\mathrm{N}}_t$ is the number of elements of $\widetilde{\mathrm{N}}_t$ (DeGroot,1970).

Given that the stochastic process on harmful event follow (3) and (4), an individual cannot make an accurate update unless he knows the exact number of people who purchased a product of the troubled company, $\widetilde{N}_t$; an official data on this number may be thought of as a formal social learning device. As is discussed in the Introduction, however, information on this number is usually not available in the market; moreover, if the troubled company would announce this number, it would be unlikely that consumers believe the information. Even in that case, consumers may gather information from various informal social learning devices; we define an informal social learning device as a tool that make information publically available with some bias and noise.

As such a device, we focus on store shelves, or, more precisely, the number of products of the troubled company that an individual sees at the stores when he buys either products of that troubled company or those of other companies that are closely substitutable to products of the troubled company in a given period. Let R_{nt} be this number, which is observed in period t by individual n. We assume that individual n uses this number to make his guess on individuals who purchase a product of the troubled company in period t, $\widetilde{N}_{nt}$, by

$$\widetilde{N}_{nt} = \kappa_n R_{nt} \tag{5}$$

It is reasonable to assume that the occurrence of a harmful event to a particular individual in a period becomes public knowledge at the end of the period; in the idiosyncratic product harm case of Aqli Foods, no poisoning incidence was reported after the first incidences in 2013. Thus, for our estimation period, we may assume that $\lambda_{mt} = 0$ before and after the incidence. Since $\sum_{m \in \tilde{N}_{nt}} \lambda_{mt} = 0$ in this case, with (5), an individual can calcurate the mean and variance of his posterior belief

$$E_{nt}[\theta] = \frac{\varphi_{n,t-1}}{\varphi_{n,t-1} + \psi_{n,t-1} + \kappa_n R_{nt}} \tag{6}$$

and

$$Var_{nt}[\theta] = \frac{\varphi_{n,t-1}(\psi_{n,t-1} + \kappa_n R_{nt})}{(\varphi_{n,t-1} + \psi_{n,t-1} + \kappa_n R_{nt})^2 (\varphi_{n,t-1} + \psi_{n,t-1} + \kappa_n R_{nt} + 1)}. \tag{7}$$

2.4. Negative Effects of Potentially Health Hazardous Products

Denote as $\mathrm{J} = \{j | j = 1, \dots, J\}$ the set of barcode-level products. All manufacturers attach a unique barcode to every distinct product; for example, a 360 ml can and a 500 ml can of one type of beer of a single company are given different barcodes even if the contents are exactly the same.

An individual receives a (negative) utility from purchasing a defective product. We assume that this disutility stems from (1) purchasing, and consuming, a defective product, which occurs when $\lambda_{nt} = 1$, (2) a risk of purchasing a defective product, and (3) the negative impression on the products of the troubled manufacturers that the individual acquires from a formal social learning device, or, namely, newspapers. That is, this negative utility for individual n in period t for product j is:

$$\Delta_{njt} = D_j \left(\omega_n \lambda_{nt} - \omega_n r_n \lambda_{nt}^2 + \nu \sum\nolimits_{\tau=0}^{t} \frac{w_\tau}{(1 + \delta_n)^{t-\tau}} \right). \tag{8}$$

In this expression, D_j is the dummy variable that assigns $D_j = 1$ if j is a product of the troubled manufacturer. Inside of the parentheses, parameter $\omega_n (< 0)$ represents consumer n's sensitivity to the health damage, $\lambda_{nt} \in \{0,1\}$, which is a random variable ex-ante; ex-post, as is noted above, $\lambda_{nt} = 0$.

Following Erdem and Keane (1996) and Ching (2010), we represent by a quadratic form of the first and second term the individual's risk attitude toward a possible health damage, where r_n $(< 1/2)$ is consumer n's risk sensitivity; the individual is risk averse if r_n is negative, risk neutral if $r_n = 0$ and risk taking if $r_n > 0$.

The third term captures the discounted sum of bad news on the troubled company carried by newspapers, a major formal social learning device. More specifically, $w_t > 0$ is the volume of bad news reported in period t by newspapers. Individuals tend to efface old news, which is captured by the discount rate δ_n (≥ 0), whose extent can differ across consumers. ν (< 0) represents the impact of such public impression on consumers' utility.

2.5. Discrete Choice Model

In describing the purchasing decision of a consumer facing a potentially health hazardous product, we add the aspect of social learning and other key variables to McFadden's discrete choice model. In addition to the expected health damage, various factors influence an individual's purchase decision. It is, among others, the prices of items that he faces, the types of goods, and personal tastes towards particular items. Let p_{njt} be the price of barcode item j individual n faces on day t, X_j be the publicly known attribute(s) of item j, and z_{nj} be individual n's personal evaluation of j that is unobservable for others and independent of the product trouble. The barcode levels of products can be categorized into many layers of different types of products. Individual n's utility when he buys j on day t is

$$U_{njt} = \alpha_{nj} p_{njt} + \beta'_{nj} X_j + z_{nj} + \Delta_{njt} + \varepsilon_{njt}, \tag{9}$$

where α_{nj} represents individual n's price sensitivity to j and β_{nj} is a vector of individual n's sensitivity to each attribute of j. ε_{njt} captures other factors that affect utility and is known by the consumer but unobservable to researchers. We assume ε_{njt} to be an i.i.d. extreme value varying across consumers, products, and time.

By using equation (5), (6), and (7), the expected utility for barcode item j, (9), can be expressed as

$$E_{nt}\left[U_{njt}\right] = \alpha_{nj} p_{njt} + \beta'_{nj} X_j + z_{nj} + D_j \left\{ \frac{\omega_n \varphi_{n,t-1}}{\varphi_{n,t-1} + \psi_{n,t-1} + \kappa_n R_{nt}} \cdot \right.$$

$$\left. \left[1 - r_n \cdot \frac{\varphi_{n,t-1} + \psi_{n,t-1} + \kappa_n R_{nt} + \varphi_{n,t-1} / (\varphi_{n,t-1} + \psi_{n,t-1} + \kappa_n R_{nt})}{\varphi_{n,t-1} + \psi_{n,t-1} + \kappa_n R_{nt} + 1} \right] + \nu \sum_{\tau=0}^{t} \frac{w_\tau}{(1+\delta_n)^{t-\tau}} \right\} + \varepsilon_{njt}, \tag{10}$$

where E_{nt} is the expected value operator conditional on the information that individual n has on day t. Note that if an individual n buys a non-troubled item, j, he faces no uncertainty so that

$$E_{nt}\left[U_{njt}\right] = U_{njt}. \tag{11}$$

The set of alternative products that a consumer can choose from may vary over time. Let J_{nt} be the set of alternative products from which consumer n choose the item to buy on date t. In summary, if consumer n buys on day t an item i, it must hold that

$$E_{nt}[U_{nit}] > E_{nt}[U_{njt}] \tag{12}$$

for all j $(\neq i)$ $\in \mathrm{J}_{nt}$.

3. Data and Modelling

As is spelled out above, our basic model permits individual heterogeneity to the maximum extent. Although *Intage SCI* data is rather big, data is still too disperse relative to the heterogeneity of individuals allowed for in our model. In our actual estimation, therefore, we condense the data, limit the heterogeneity of the model, and yet maintain the versatility of the data and the generality of the model to characterize a social learning process.

3.1. Aggregation and Grouping of Barcode-wise Products

A. Daily Data to Weekly Data: Although the *Intage SCI* data is daily panel data, we aggregate it into weekly data. This is mainly because of the actual purchasing patterns of individuals regarding frozen foods. Although there is an enormous diversity, the number of items each individual bought per month, in our dataset, was about 4.12 on average. Using daily data may, therefore, not only generate unnecessarily sparse panels, but also capture complex miscellaneous behaviors of individuals such as stock decisions, which are out of our research purpose and our model. On the other hand, using monthly data may fail to capture delicate changes in people's behaviors facing the ever-changing situations over the incident. Unless otherwise specified, therefore, we regard variable t, by which we indicated day in the previous section, as indicating week.

B. Grouping of Barcode-level Products: Some frozen foods are similar (closely substitutable) one another. But others are not. In order to capture this fact, we introduce a layered structure of goods. As is explained below, this structure consists of barcode items, "groups" of barcode items, and "categories" of foods. This is not only to reduce the number of parameters to be estimated but also to control each individual's preference over different categories and groups of items and to highlight the effect of social learning in individual purchase decisions. At the same time, by adopting the layered structure, we hope to capture who might become early purchasers of troubled products.

In the poisoning incident, Aqli Foods recalled six categories of frozen foods: (Category 1) croquettes, which are popular appetizer/side-dish for children's lunchbox, (Category 2) gratins/dorias/lasagnas, which are of dinner size, (Category 3) dinner-size pizzas, (Category 4) small-size side dishes, which are mainly used for children's lunch box, (Category 5) pies and pie sheets, and (Category 6) hotcakes and pancakes. We capture the preference orderings over these categories of foods by assuming that for each category, c, each individual's sensitivity to the price and the observed characteristics of each barcode item do not vary across items, i.e., for each c, that there are constant α_{nc} and β_{nc} such that

$$\alpha_{nj} = \alpha_{nc} \text{ and } \beta_{nj} = \beta_{nc} \text{ for all } j \in J_c \tag{13}$$

where J_c represents the barcode items belonging to category c.

In order to detect the effect of price and attributes separately from that of product quality, we combine a particular manufacturer's particular type of barcode items into one group if those items provide similar quality. If, for example, one company may sell "Frozen Margherita Pizza 300g" and "Frozen Margherita Pizza 500g," we classify these items into one group, even if their quantities differ (so that different barcodes are given). If that company sells "Frozen Margherita Pizza 300g" and "Frozen Cheese Pizza 300g" as well, they are classified into different groups. In equation (9) and (10), the quality that an individual perceives for a particular group of barcode items is captured by z_{ng}, where

$$z_{nj} = z_{ng} \text{ for all } j \in J^g \tag{14}$$

where J^g represents the barcode items belonging to group g.

By using the above aggregations and groupings, we can rewrite equation (9) and (10) as

$$U_{ncjt} = \alpha_{nc} p_{njt} + \beta'_{nc} X_j + z_{ng} + D_j \left(\omega_n \lambda_{nt} - \omega_n r_n \lambda_{nt}^2 + \nu \sum_{\tau=0}^{t} \frac{w_\tau}{(1+\delta_n)^{t-\tau}} \right) + \varepsilon_{ncjt}, \tag{9'}$$

and

$$E_{nt}\left[U_{ncjt}\right] = \alpha_{nc} p_{njt} + \beta'_{nc} X_j + z_{ng} + D_j \left\{ \frac{\omega_n \varphi_{n,t-1}}{\varphi_{n,t-1} + \psi_{n,t-1} + \kappa_n R_{nt}} \cdot \right.$$

$$\left. \left[1 - r_n \cdot \frac{\varphi_{n,t-1} + \psi_{n,t-1} + \kappa_n R_{nt} + \varphi_{n,t-1} / (\varphi_{n,t-1} + \psi_{n,t-1} + \kappa_n R_{nt})}{\varphi_{n,t-1} + \psi_{n,t-1} + \kappa_n R_{nt} + 1} \right] + \nu \sum_{\tau=0}^{t} \frac{w_\tau}{(1+\delta_n)^{t-\tau}} \right\} + \varepsilon_{ncjt}. \tag{10'}$$

3.2. Sample Period and Sample Individuals

We use the purchase records of the relevant frozen foods for one and half year from the first week of April 2014 to the last week of December 2015 to estimate the model. We choose this period to focus on the sales recovery process of Aqli Foods in its post-crisis period. We include the data before August 2014, when the Gunma factory partially restarted its production, since Aqli Foods had kept its production in the other factory (Yubari Factory) before August. But we exclude the period before April 2014 because, at least in our dataset, Aqli products were sold at only a tiny fraction of retailers during the period and it was after the beginning of April when Maruha Nichiro announced the efforts to seek concrete measures to strengthen food defense that the products were sold at broader range of retailers.

In the *Intage SCI* data, there are many individuals who rarely buy frozen foods. In order to exclude those individuals, we first select the "most active" 100 survey participants. That is to say, we select the survey participants who participated in the survey one year before and two years after the poisoning incident (January 2013 - December 2015) and visited during the estimation period (April 2014 - December 2015) at least once a store that sold at least one of the 148 barcode items, as will be explained below, of Aqli Foods that we use in the estimation. We line up those participants in the order of the number of purchases of a frozen food product over the estimation period and selected to the top 100. By doing so, we can eliminate the consumers who did not purchase Aqli Foods's products because the retail stores he or she visited did not sell them so that he or she did not have an opportunity to purchase them. Table 1 reports summary statistics on some major attributes of the 100 consumers. The majority are married females, 30–59 years old. The job statuses are mainly part time (46%), regular employee or public service (17%), and homemakers (19%).

3.3. Sample Categories for Purchase Decision

As is noted above, our data consists of 680,000 barcode items. In order to focus on the products affected by the Aqli poisoning incident and to render our model tractable, we select only 148 barcode items in the top three categories of frozen foods, denote as J_c^* $(c = 1, 2, 3)$; this selection is explained below in more detail. Denote as $\mathrm{J}_{nct}^* \subseteq \mathrm{J}_c^*$ the set of the items in category c that are sold at the stores that individual n visits at t. Following the standard literature of consumer learning using a discrete choice model (e.g. Erdem and Keane, 1996; Erdem, 1998; Ackerberg, 2003), we assume that a survey participant compares all items sold at the stores he visits and receives utility U_{ncjt} if he buys a primary item $j \in \mathrm{J}_{nct}^*$ from category c and a reservation utility U_{nc0t} if he buys a non-primary item $j \notin \mathrm{J}_{nct}^*$ including products other than frozen foods.[7]

[7] It is possible that a survey participant visited the store but decided not to buy any, but since we cannot distinguish, only from our data, such a case from the case in which he did not even visit the store for reasons other than daily consumption decisions, we do not use non-purchase cases in the estimation.

We assume that if an individual buys a non-primary item, his utility is at a reservation level, i.e.,

$$U_{nc0t} = \chi_n + \varepsilon_{nc0t}, \tag{15}$$

where χ_n is the mean utility level that individual n acquires by purchasing a non-primary item and ε_{nc0t} represents a fluctuation from it, both of which are observable only by him.

The top three categories of frozen foods that we include in our estimation are: Category 1 (croquettes), Category 2 (gratins/dorias/lasagnas), and Category 3 (dinner-size pizzas). Note that the highest monthly sales during the period from April 2012 to March 2016 are about 1.8M yen for Category 1, 1.4M for Category 2, 0.8M for Category 3, 0.8M for Category 4, 0.4M for Category 5, and less than 0.1M for Category 6. Although the volume of sale for Category 4 (small side dishes) is fairly large, we do not include Category 4 items in our estimation. This is because it is practically difficult to obtain common credible estimates for the price sensitivity and/or weight sensitivity for Category 4, which consists of much more diverse types of foods than the first three categories.

The number of barcode items belonging to the three categories sold over the period from a year before the detection of the incident to two years after it (January 2013 - December 2015) is 740; more precisely, Categories 1, 2, and 3 consists of, respectively, 334, 281, and 125 items.[8] But because the inclusion of less popular items render the estimation unstable, we first exclude the non-Aqli Foods' products of which the monthly share in each category was not ranked within the 15 largest over the above period and treat them as non-primary items.[9] By checking the packages and descriptions of products, we classify the remaining barcode items in each category into groups; they consists of 49, 38, and 50 groups, respectively, in Categories 1, 2, and 3. We then calculate sales shares of each group in the total sales of all items in each category sold during the estimation period (April 2014 - December 2015) including the ones that we excluded from the primary items in the above procedure, and treat the groups with above 1.5% shares as independent groups in the estimation. Also, we group the Aqli Foods' products with less than 1.5% share as one unified group (hereafter, "other Aqli products"), and the non-Aqli Foods' products with less than 1.5% share as another unified group (hereafter, "other non-Aqli products").[10] See Figure 1, which shows each manufacturer's sales in the three categories. The colors and patterns in each graph represent individual manufactures and the areas surrounded by the lines represent individual product groups. Finally, we exclude the groups that the 100 participants did not purchased during the estimation period, and we have 33 groups and 148 barcodes left as the primary items (10, 14, and 9 groups, and

[8] We exclude 3 items from Category 3 here since we cannot find any weight data for these items.

[9] This accounts for 270, 188, and 31 items, respectively, in Category 1, 2, and 3.

[10] The items that were sold only at online stores or by mail-orders and the items that were not sold during the estimation period are excluded from "other Aqli products" and "other non-Aqli products."

41, 68, and 39 barcodes, respectively, in Category 1, 2, and 3), of which 10 groups and 52 barcodes are Aqli Foods' products (2, 4, and 4 groups, and 6, 30, and 16 barcodes, respectively, in Categories 1, 2, and 3).

3.4. Data on Choice Sets, Prices, and Observable and Unobservable Characteristics

A. Choice Sets and Price Data: In our estimation, we need data on the lists of items that an individual survey participant considered but did not buy as well as those that he bought. Data on the items that a survey participant actually purchased can immediately be collected from our scanner panel data. If an individual bought more than one unit of an item in a week, we use the weekly average of the prices he actually paid during that week.

Data on an item that a survey participant considered but did not purchase is far more difficult to obtain. First, even if a purchase decision is limited to the 148 primary items in the three categories of frozen foods, it is hard to imagine that an individual compares the prices and other characteristics of all items that are available at a store. An individual would, however, compare explicitly those of at least several similar items in front of his eyes before selecting one. Moreover, back of his mind, he would have usual prices or other characteristics for the items that he does not explicitly but implicitly consider. For example, if he would consider buying a frozen pizza at a store, he would implicitly consider the prices and other characteristics of several different pizzas at other stores before selecting one. Second, the data on items that a participant did not purchase is not directly available in the *Intage* scanner panel data, which does not compile the barcode items that each store carries; scanner panel data is data on what, where, and when each consumer buys.

From the *Intage* data, however, we may approximately know the choice sets of individuals by using the purchase data. The idea is as follows. If, for example, a particular individual does not buy a particular item at a particular store, some others are likely to buy it at that store; otherwise the store would not carry it. If, therefore, all Japanese people were *Intage* respondents, it would be possible fairly accurately to capture which stores carry the item. With the 50,000 *Intage* respondents, however, this would unlikely be the case. We overcome this problem by relying on the fact that a large number of stores at which the respondents make their purchases are members of a chain store group or a large retail company. With this consideration, we assume the following.

Assumption 1: *If a particular product is sold at a store in a particular month, it is sold, in that month, at all stores in the same prefecture that belong to the same retail company or chain group.*

Specifically, we take the following procedure. We first collect, for each retail company operating in a prefecture and for each week, the list of primary items and their prices that survey participants purchased and paid in that week at the stores operated in that prefecture under that company, calculate the average price for each item, and use this list of items and their average prices as the items and their prices in the week at any store operated in the prefecture by the company. There are weeks in which no survey participants bought the item at any store operated by the company, but some participants bought it in other week in the same month. For such a week, we collect the monthly item list and calculate the monthly average prices in much the same way and use them as the items and their prices in that week at any store operated in the prefecture by that company. When constructing these lists of items and the average prices, we use the purchase records of not only the 100 participants we use in the estimation but about 50,000 participants included in the original dataset, so that we can construct much richer sets of prices and products for each store for each week.

Once the above dataset is built, for each survey participant, n, we collect all the stores at which he made at least one purchase of any frozen foods not limited to the primary items during a given week, t, and make a list of primary items sold at those stores and their prices for each category. We identify sum of this set of barcode items and the items that the participant purchased, denoted as J^*_{nct}, as participant n's choice set of category c in week.

B. Data on Observable (Objective) Attributes: Such items as cars and PCs are associated with a large number of attributes. The attributes with which foods (or frozen foods) are associated are relatively limited. Among them, perhaps important are weight/calorie and materials/ingredients. Of those, this study focuses on content weight. For each barcode item, we check the product description, obtain weights and use them as data for X_j. Some of those weights are available in the *Intage* data. For those that are not specified in the *Intage* data, we obtain relevant data from the Internet sites of manufacturers or other webpages such as online shopping site or customer review sites.

C. Unobservable (Subjective) Attributes: No data is available for unobservable attributes. By the grouping discussed above, we estimate z_{ng} for each group of items.

3.5. Data on Signals from the Informal Social Learning Device

As is noted above, if the number of consumers who buy Aqli products in each period (a formal social learning device), $\widetilde{N}_t$, were publically available, each individual could update his belief on the safety of Aqli products by using equations (3) and (4). We assume that, without such a formal device, consumers adopt an informal social learning device in the form of store shelves, more precisely, the number of store shelves which a consumer sees, and registers to have seen, sell products of the troubled company in each period, R_{nt}; with R_{nt}, his belief is updated by equations (3), (4) and (5). Specifically, we take the following procedure in constructing the data of R_{nt}. For each day, for each survey participant, and for each store that he visited, we count the number of categories of the six from which he purchased some items, not limited to Aqli products, and in which the store sold Aqli products. The list of items that the store sold on that day is constructed through the procedure explained in the previous section. By doing so, we can detect the opportunities in which a participant is likely to have observed Aqli products when he was considering which items to buy. Then we sum up these numbers across all the stores he visited during that week and let R_{nt} denote this sum.

3.6. Volume of Bad News Obtained through a Formal Social Leaning Device

.

Recall that w_t is the volume of bad news reported in period t. We compile this data in the following way. We collect articles related to the incident from the four largest newspapers in Japan, Asahi, Nikkei, Yomiuri, and Mainichi. We ask seven monitors[11] to read article sets of each newspaper in chronological order and rate each article "1" if he or she feels the credibility of Aqli Foods has declined, "−1" if increased, and "0" if unchanged. The names of the newspapers are concealed from the monitors. For the second, third, and fourth sets, the monitors answer the extent to which they feel affected by the contents of the previous set, which we call "affected rate." For each of the four newspapers, we construct weekly averages for the seven monitors by using the numbers of letters of each article as weights and by discounting each monitor by his or her affected rate. Then we construct weekly weighted-averages for the four newspapers based on their circulations.

4. Estimation Method

The above model of discrete choice, $(10')$ and (15), has a number of parameters, concerning individual tastes: They are α_{nc}, β_{nc}, z_{ng}, χ_n, ω_n, κ_n, r_n, and δ_n. Among them, of particular importance are κ_n, capturing the reliance of individual n on informal social learning device, δ_n, capturing how quickly an individual effaces bad memories obtained through a formal social learning device, and r_n, capturing risk aversion. Following the literature of consumer learning (e.g. Erdem and Keane, 1996; Erdem, 1998; Erdem, Zhao, and Valenzuela, 2004; Iyengar, Ansari, and Gupta, 2007; Zhao, Zhao, and Helsen, 2011), we assume that the rest of parameters obey normal distributions with unknown mean and variance μ_π and σ_π^2 (hyperparameters). That is,

$$\pi_n \sim N(\mu_\pi, \sigma_\pi^2), \qquad \pi = \alpha_c, \beta_c, z_g, \chi, \omega_n \tag{16}$$

where we assume $\sigma_{\pi_n} > 0$ (which implies all the standard deviations are positive); $\mu_{\alpha_{nc}} < 0$ and $\alpha_{nc} < 0$ (which implies that a price increase reduces utility); $\mu_{\beta_{nc}} > 0$ and $-\infty \quad < \beta_{nc} < \infty$ (which implies that on

[11] Of the seven monitors, three are college students; 4 are part-time employees.

average individuals like a more content but that there can be someone who may feel otherwise); $\mu_{z_{ng}} > 0$ and $z_{ng} > 0$ (which implies that individuals have positive perceived qualities); $\mu_{\chi_n} = 1$ and $\chi_n > 0$ (the first expression of which implies the normalization of reservation utility and the second, the assumption that everyone receives a positive utility when he buys a non-primary item); $\mu_\omega < 0$ and $\omega < 0$ (which implies that the realization of product harm reduces the utility); and $\mu_r < 0$ and $-\infty < r < \infty$ (which implies that on average individuals are risk averse but that there can be some risk lovers). We need the normalization of reservation utility $\mu_{\chi_n} = 1$ since the absolute values of the mean utility levels of non-primary items, χ_n have no meaning. The absolute values of perceived qualities, z_{ng}, are also pinned down by this normalization through comparisons between primary and non-primary items. Moreover, we assume that heterogeneity in individual preferences are fully controlled by these parameters so that the residuals in equations (10′) and (15), ε_{ncjt} and ε_{nc0t}, $j \in J \cup \{0\}$, follows i.i.d. Gumbel and type 1 extreme value.

For informal learning coefficient κ_n and the discount factor of past bad news, δ_n, we need $\kappa_n \geq 1$ and $\delta_n \geq 0$ from the nature of these variables, but we do not impose further restrictions on their distribution over the population. The initial values of ψ_{n0} are set at $\psi_{n0} = 1$ and those of $E_{n0}[\theta]$ at $E_{n0}[\theta] = 0.01$ for all consumers. The initial values of φ_{n0} are then given by $\varphi_{n0} = E_{n0}[\theta]/1 - E_{n0}[\theta]$.

By the construction of our model discussed in the previous section, it holds that $E_{nt}[U_{ncit}] \geq E_{nt}\left[U_{ncjt}\ \right]$ for any j $(\neq i)$ $\in \mathrm{J}^*_{nct}$ and $E_{nt}[U_{ncit}\ ] \geq U_{nc0t}$ if individual n buys a primary item $i \in \mathrm{J}^*_{nct}$, and that $U_{nc0t} \geq E_{nt}\left[U_{ncjt}\right]$ for any $j \in \mathrm{J}^*_{nct}$ if individual n buys a non-primary item . We estimate the model by the hierarchical Bayesian estimation with MCMC simulations.

For the items belonging to "other Aqli products" and "other non-Aqli products," we use the price data, p_{njt}, which varies across individuals, barcode items, and weeks, and the weight data, X_j, which varies across barcode items. But we estimate z_{ng} for those items by assuming that an individual perceives the similar quality for all items belonging to the same unified group.

5. Estimated Hyperparameters

Before examining the role of social learning, it is desirable to doublecheck if the estimated model is economically reasonable. Tables 2 and 3, respectively, report the means, their standard errors, and standard deviations of estimated hyperparameters, (μ_π, σ_π), respectively, for $\pi = \alpha_c$, β_c, χ, ω , and r, and the parameter χ, and those of hyperparameters $(\mu_{z_g}, \sigma_{z_g})$ for $g = 1, \ldots, 33$. Figure 2 shows the distributions of the individual-level parameters, α_{nc}, β_{nc}, κ_n, ω_n, r_n, and δ_n, all of which are located over the week that each individual first purchased Aqli products during the estimation period. With regards to α_{nc} and β_{nc}, the vertical values represent the individual's average of the three product categories for each individual (let $\bar{\alpha}_n$ and $\bar{\beta}_n$ denote these average). The individuals who did not buy Aqli products are shown as red dots at the right edge of each scatter diagram.

As Table 2 shows, the hyperparameters on the basic preference coefficients, μ_{α_c} and μ_{β_c} are significantly different from zero. Since $\mu_{\alpha_c} < 0$ and $\mu_{\beta_c} > 0$, the individual utility is, on average, likely to be negatively related to the price and positively to the weight in all three categories of frozen foods: (1) croquettes, (2) gratins/dorias/lasagnas, and (3) dinner-size pizzas. Compared to the price sensitivity, the sensitivity to a change in weight shows relatively large variation across individuals. This is true for all three categories but, especially, Category 1 (croquettes) shows larger variations; the mean hyper standard deviation ($\sigma_{\beta_1} = 0.008$) is about twelve times larger than the mean hyper mean ($\mu_{\beta_1} = 0.0007$). In fact, as shown in Figure 2, the individual-level parameters, β_{nc}, are negative for some participants. This indicates some consumers prefer smaller sizes, presumably depending on family size or food habit, although the majority prefer larger ones. The reservation utility that an individual receives when he buys an item other than the primary products also varies fairly widely across individual, $\sigma_\chi = 6.307$. This makes sense because peoples' demands for those products reflect a wide variety of factors.

The health-damage sensitivity, ω_n, should be negative for its nature. As shown in Figure 2, the individual-level parameters are all estimated to be negative, although the statistical significance of the hyper mean, μ_ω, is slightly week. The risk attitude, r_n, are all estimated to be negative, indicating that all participants are risk averse. In addition to these hyperparameters, we present estimation results on ν, the sensitivity to bad news. This parameter is not a hyperparameter but a parameter capturing the sensitivity to bad news, which is common for all individuals. It is significantly negative, indicating that the negative public impressions on Aqli Foods have negative impact on consumers' purchase decisions.

Table 3 shows the hyper means and hyper standard deviations of the perceived qualities of each primary items. People's perception on the qualities of Aqli products, aside from the poisoning incident, are on average not bad compared to the products of competing companies in all three categories. Especially, the higher values in Category 3 (dinner-size pizzas) indicate the strong popularity of Aqli products, which coincides with the large sales share of Aqli Foods in this category even after the incident, as shown in Figure 1.

6. Determinants of a Recovery from an Idiosyncratic Negative Shock

In this section, we examine the role of an informal social learning device in the process of a recovery from an idiosyncratic negative shock. For this purpose, we compare the effect of an informal social learning device with those of other factors that may contribute to the recovery process, in particular, that of a formal macro device for social learning, or, more specifically, of newspaper articles on the negative shock.

6.1. Role of an Informal Social Learning Device (Store Shelves)

Parameter κ_n represents the extent to which individual n relies on store shelves as an informal device for social learning. For each estimated κ_n, we may calculate a Bayesian confidence interval; in Figure 3.A, we present the 95% confidence intervals for all κ_n, $n = 1, \dots, 100$. All estimated distribution of κ_n are significantly away from 0. This suggests that store shelves serve as a social learning device.

Finding 1: *For every individual, the store shelves that he sees serve as an informal social learning device.*

As Figure 2.A shows, while the variance of each κ_n is fairly large, κ_n do not vary across individuals.

6.2. Role of the Formal Social Learning Device (Newspaper Articles)

Parameter δ_n represents how quickly an individual effaces bad memories that he obtained through a formal social learning device, or, in our model, newspaper reports on the idiosyncratic negative shock to Aqli products. Figure 3.B presents the estimated value and its 95% Bayesian confidence interval for each estimated δ_n, $n = 1, \dots, 100$. As the figure show, all estimated δ_n are significantly away from zero. This suggests that the effacing of past bad news contributes to the recovery from an idiosyncratic negative shock.

Finding 2: *For every individual, the effacing of past bad news in a formal social learning device contributes to the recovery process from an idiosyncratic negative shock.*

6.3. The Role of Individual Diversity

In order to illustrate the coexistence of returners and non-returners, Figure 3 plots the vector of an individual's estimated parameter value and the week in which that individual started purchasing Aqli products for $\bar{\alpha}_n$, $\bar{\beta}_n$, κ_n, ω_n, r_n, and δ_n. Because we use data through the first 91 weeks from the beginning of April, 2014, the red points on the vertical line at the 92nd week indicates those who did not purchase Aqli products for the first 91 weeks. In what follows, we call those individuals non-returners and those who started to buy Aqli products within the first 91 weeks returners.

It is natural to assume that the more demand for a particular product, the more likely a store carries that product. If so, the existence of returners contributes to the recovery process for the products of Aqli Foods, which suffered from selling a product with an idiosyncratic defect. In this sense, the coexistence of returners and non-returners shows that the diversity reflected in the separation of the two groups, individual activities, individual preferences and market conditions plays key roles in the informal social learning towards a recovery from an idiosyncratic product defect.

Finding 3: *Diversity that individuals face plays a key role in informal social learning during a recovery process from an idiosyncratic negative shock.*

Figure 2.C suggests that the higher an individual's reliance on the informal social learning device (or the larger κ_n), the more likely that an individual starts buying the troubled good. This is clearly shown by a comparison between returners and non-returners. As Table 4 (C) shows, the average κ_n for non-returners (30 individuals) is 471.0 whereas that for returners (70 individuals) is 521.1. The highest κ_n for non-returners is $\kappa_n = 494.6$. In contrast, 76 percent of the 70 returners have a value of κ_n above that maximum value, $\kappa_n > 494.6$.

Moreover, Figure 2.F suggests that the faster an individual effaces past bad memories (or the larger δ_n), the faster that an individual starts buying the troubled good. As Table 4 (F) shows, the average δ_n for non-returners (30 individuals) is 0.097 whereas that for returners (70 individuals) is 0.102. The highest δ_n for non-returners is $\delta_n = 0.102$. In contrast, 53 percent of the 70 returners has a value of δ_n above that maximum value, $\delta_n > 0.102$. These facts suggest that the recovery process from an idiosyncratic negative shock to Aqli products is supported by the diversity of individual attitudes in using store shelves as an informal social learning device.

Table 4 reports the mean and standard deviation, over participants, of the mean of each of individual parameters for the groups of returners and non-returners. For all parameters but β_n, the standard deviation is higher for the returner group than for the non-returner group. It is also clear from Figure 2(C) and (D) that κ_n and δ_n have higher mean and larger variance for the returner group than for the non-returner group. All these findings further support that the diversity of individuals are important for the recovery process from an idiosyncratic negative shock.

An interesting fact captured in Figure 2 is that around the fortieth week (December 29, 2014 – January 5, 2015), for some parameters, especially $\bar{\alpha}_n$, ω_n, and δ_n, the returners may be divided into two groups: early returners and late returners. In fact, as Table 4 shows, the early returners, on average, have a lower α_n, a higher β_n, a higher κ_n, a higher ω_n, and a higher δ_n than either the non-returners or the later returners. These facts all make an intuitive sense. That is, (1) the less price sensitive, (2) the more content weight sensitive, (3) the more heavily reliance on the informal social learning device, (4) the less risk sensitive, and/or (5) the more readily effacing bad memories, the more likely an individual to return to goods supplied by a company that experienced an idiosyncratic negative shock.

6.4. Effacing of Past Bad News

Our analysis captures the way in which individuals efface bad news provided by newspapers, a formal social learning device in a broad sense; individuals tend to efface bad news drastically around the thirtieth to fortieth week. In order to check this, Figure 4 shows the time transition in the average amount of bad memories, $\sum_{\tau=0}^{t} w_t /(1+\delta)^{t-\tau}$, with the average of individual discounting rates, $\delta = \sum_n \delta_n /100 \approx 0.101$. The average amount of bad memories makes a sharp rise after the first announcement of the pesticide detection on December 29, 2013. Although it starts declining in the middle of February as the number of articles decreases after the arrest of the suspect on January 25, 2014, it increases again responding to the interim report of the independent examination committee published on April 30, 2014 that revealed the poor risk management and insufficient food defense of Aqli Foods and Maruha Nichiro during the crisis. The average amount of bad memories thereafter trended toward improvement, which shows even acceleration when the president of Maruha Nichiro published the measures to prevent recurrence in response to the final report of the committee on May 29, 2014 and when it was reported that the Gunma factory restarted its production on August 1, 2014. Around the end of September 2014, it declines to the range below 10, which represents a level of 1/20 of the peak, and then converged to zero in a gentler slope.

References

Ackerberg, D. A. (2003). "Advertising, Learning, and Consumer Choice in Experience Good Markets: An Empirical Examination," *International Economy Review* (Philadelphia) 44, 1007–1040.

Akçura, M. T., F. F. Gönül, and E. Petrova. (2004). "Consumer Learning and Brand Valuation: An Application on Over-the-Counter Drugs," *Marketing Science* 23, 156–169.

Akhtar, N. and A. Mian. (2018). "Nonparametric Coupled Bayesian Dictionary and Classifier Learning for Hyperspectral Classification," *IEEE Transactions on Neural Networks and Learning Systems* 29, 4038–4050.

Associated Press (2014). "Frozen food tainted with pesticides sickens 350 people in Japan," *New York Daily News*, Jan. 7, 2014.

Besbes, O., and M. Scarsini. (2018). "On information distortions in online ratings," *Operating Research*, 66, 597–610.

Banerjee, A., (1992). "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107-3, 797-817.

Banerjee, A., and D. Fudenberg (2004). "Word-of-mouth Learning," *Games and Economic Behavior* 46, 1-22.

Bikhchandani, S., D. Hirshleifer and I. Welch, (1992). "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy* 100-5, 992-1026.

Chamley, P. C. (2003). *Rational Herds: Economic Models of Social Learning*, New York: Cambridge University Press.

Chen, B., G. Polatkan, G. Sapiro, D. Blei, D. Dunson, and L. Carin. (2013). "Deep learning with hierarchical convolutional factor analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35, 1887–1901.

Ching, A. T. (2010). "Consumer Learning and Heterogeneity: Dynamics of Demand for Prescription Drugs after Patent Expiration," *International Journal of Industrial Organization* 28, 619–638.

Davis, A., V. Gaur, and D. Kim. (2020), "Consumer Learning from Own Experience and Social Information: An Experimental Study," SSRN, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3208758

DeGroot, M. H. (1970). *Optimal Statistical Decisions*. Hoboken, NJ: John Wiley & Sons, Inc.

Erdem, T. (1998). "Empirical Analysis of Umbrella Branding," *Journal of Marketing Research* 35, 339–351.

Erdem, T. and M. P. Keane. (1996). "Decision-Making Under Uncertainty: Capturing Dynamic Brand Choice Processes in Turbulent Consumer Goods Markets," *Marketing Science* 15, 1–20.

Erdem, T., Y. Zhao, and A. Valenzuela. (2004). "Performance of Store Brands: A Cross-Country Analysis of Consumer Store-Brand Preferences, Perceptions, and Risk," *Journal of Marketing Research* 41, 86–100.

Herrera, H., and J. Hörner. (2013). "Biased social learning," *Games and Economic Behavior*, 80, 131–146.

Honryo, T. and M. Yano. (2020). "Idiosyncratic information and vague communication," *American Political Science Review*, 2020, 1-14.

Hoppitt, W. and K. N. Laland. (2013). *Social Learning: An Introduction to Mechanisms, Methods, and Models*. Princeton, NJ: Princeton University Press.

Israel, M. (2005). “Services as Experience Goods: An Empirical Examination of Consumer Learning in Automobile Insurance,” *American Economic Review* 95, 1444–1463.

Iyengar, R., A. Ansari, and S. Gupta. (2007). “A Model of Consumer Learning for Service Quality and Usage,” *Journal of Marketing Research* 44, 529–544.

Kyodo (2014). “Suspect told police how to poison food,” *Japan Times*, Jan. 27, 2014.

Lee, H. J. and K. S. Hong. (2016). “Class-specific mid-level feature learning with the Discriminative Group-wise Beta-Bernoulli process restricted Boltzmann machines,” *Pattern Recognition Letter* 80, 8–14.

McFadden, Daniel. (1974). “Conditional Logit Analysis of Qualitative Choice Behavior,” Zarembka ed., *Frontiers in Econometrics*, 106-142, Academic Press, New York.

Meager, Rachael. 2016. “Aggregating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature.” https://economics.mit.edu/files/12292.

Meager, Rachael. 2019. “Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments,” *American Economic Journal: Applied Economics*. https://doi.org/10.1257/app.20170299

Narayanan, S. and P. Manchanda. (2009). “Heterogeneous Learning and the Targeting of Marketing Communication for New Products,” *Marketing Science* 28, 424–441.

Roberts, J. H. and G. L. Urban. (1988). “Modeling Multiattribute Utility, Risk, and Belief Dynamics for New Consumer Durable Brand Choice,” *Management Science* 34, 167–185.

Shen, Z. and P. M. Djurić. (2014). “Distributed Bayesian learning with a Bernoulli model,” *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 5482–5486.

Smith, L., and P. Sørensen. (2000). “Pathological outcomes of observational learning,” *Econometrica*, 68, 371–398.

Tatusian, T. (2014). “Japan copes with tainted food outbreak,” *Los Angeles Times*, Jan. 7, 2014.

Tomlin, B. (2009). “Impact of supply learning when suppliers are unreliable,” *Manufacturing and Service Operations Management* 11, 192–209.

Train, K. (2003). *Discrete Choice Methods with Simulation*, Cambridge University Press, Cambridge.

Viet Nam News (2014). “Food poisoning case rocks Japan,” *Viet Nam News*, Jan. 9, 2014.

Vives, X. (1993). “How fast do rational agents learn,” *Review of Economic Studies*, 60, 249–282.

Zhao, X, and A. H. Sayed. (2015). “Asynchronous Adaptation and Learning Over Networks,” *IEEE Transactions on Signal Processing* 63, 811–842.

Zhao, Y., S. Yang, V. Narayan, and Y. Zhao, (2013). “Modeling Consumer Learning from Online Product Reviews.” *Marketing Science*. 32, 153–169.

Zhao, Y., Y. Zhao, and K. Helsen. (2011). "Consumer Learning in a Turbulent Market Environment: Modeling Consumer Choice Dynamics after a Product-Harm Crisis," *Journal of Marketing Research* 48, 255–267.

Table 1

N=100

Age	–20	20–29	30–39	40–49	50–59	60–
	0	3	22	44	27	4
Sex	Female	Male				
	81	19				
Marital status	Unmarried	Married				
	15	81				
Job status	Part time	Student	Self-employed	Regular employee/ Public service	Homemaker	Dispatched employee/ Contract worker
	46	1	2	17	19	12
	Other job	Unemployed				
	1	2				
Household income (10,000 Yen)	−399	400−549	550−699	700−899	900−	
	28	22	23	18	9	
Area	Hokkaido	Tohoku	Kitakanto	Capital area	Hokuriku	Tokai
	4	3	3	20	9	19
	Keihanshin	Chugoku	Shikoku	Kyushu		
	14	7	3	18		

Table 2

	Mean	SE of mean	Standard deviation		Mean	SE of mean	Standard deviation
Price sensitivity							
μ_{α_1}	-0.02307	0.00011	0.00204	σ_{α_1}	0.01617	0.00008	0.00169
μ_{α_2}	-0.01795	0.00025	0.00217	σ_{α_2}	0.00809	0.00017	0.00173
μ_{α_3}	-0.03074	0.00200	0.00564	σ_{α_3}	0.01469	0.00111	0.00360
Weight sensitivity							
μ_{β_1}	0.00070	0.00001	0.00060	σ_{β_1}	0.00819	0.00014	0.00185
μ_{β_2}	0.00123	0.00005	0.00092	σ_{β_2}	0.00717	0.00007	0.00118
μ_{β_3}	0.00581	0.00045	0.00342	σ_{β_3}	0.00808	0.00054	0.00347
Outside alternative utility							
μ_{χ}	1	Fix	Fix	σ_{χ}	6.30707	0.14527	0.73546
Health-damage sensitivity							
μ_{ω}	-300.84	178.39	317.09	σ_{ω}	60.26	20.62	74.19
Risk attitude							
μ_{r}	-713.71	494.77	1,758.77	σ_{r}	145.55	100.48	403.93
News sensitivity							
ν	-0.00295	0.00012	0.00281				

Table 3

		Mean	SE of mean	Standard deviation			Mean	SE of mean	Standard deviation
Category 1 (croquettes)									
Primary Aqli	μ_{z_1}	7.136	0.162	0.697		σ_{z_1}	1.216	0.029	0.309
Other Aqli	μ_{z_2}	6.060	0.139	0.777		σ_{z_2}	2.053	0.060	0.413
	μ_{z_3}	7.452	0.160	0.653		σ_{z_3}	1.700	0.008	0.195
	μ_{z_4}	6.487	0.159	0.668		σ_{z_4}	1.735	0.012	0.232
	μ_{z_5}	7.430	0.160	0.668		σ_{z_5}	2.136	0.012	0.247
	μ_{z_6}	6.523	0.157	0.673		σ_{z_6}	1.417	0.017	0.236
	μ_{z_7}	7.548	0.157	0.645		σ_{z_7}	1.218	0.013	0.212
	μ_{z_8}	6.297	0.155	0.701		σ_{z_8}	1.551	0.028	0.309
	μ_{z_9}	4.783	0.109	0.612		σ_{z_9}	2.023	0.055	0.347
Other non-Aqli	$\mu_{z_{10}}$	6.650	0.158	0.642		$\sigma_{z_{10}}$	1.011	0.015	0.213
Category 2 (gratins/dorias/lasagnas)									
Primary Aqli	$\mu_{z_{11}}$	5.517	0.153	0.653		$\sigma_{z_{11}}$	1.116	0.189	0.487
Primary Aqli	$\mu_{z_{12}}$	5.233	0.085	0.599		$\sigma_{z_{12}}$	1.535	0.112	0.509
Primary Aqli	$\mu_{z_{13}}$	5.788	0.099	0.595		$\sigma_{z_{13}}$	0.925	0.184	0.584
Other Aqli	$\mu_{z_{14}}$	5.465	0.154	0.661		$\sigma_{z_{14}}$	0.956	0.142	0.439
	$\mu_{z_{15}}$	4.989	0.077	0.522		$\sigma_{z_{15}}$	1.923	0.023	0.310
	$\mu_{z_{16}}$	4.628	0.064	0.513		$\sigma_{z_{16}}$	1.677	0.025	0.300
	$\mu_{z_{17}}$	5.312	0.106	0.603		$\sigma_{z_{17}}$	1.317	0.052	0.367
	$\mu_{z_{18}}$	2.778	0.079	0.471		$\sigma_{z_{18}}$	1.160	0.032	0.266
	$\mu_{z_{19}}$	5.098	0.058	0.600		$\sigma_{z_{19}}$	1.560	0.073	0.417
	$\mu_{z_{20}}$	4.279	0.074	0.468		$\sigma_{z_{20}}$	1.093	0.078	0.361
	$\mu_{z_{21}}$	4.290	0.078	0.530		$\sigma_{z_{21}}$	1.081	0.052	0.346
	$\mu_{z_{22}}$	4.219	0.069	0.501		$\sigma_{z_{22}}$	1.546	0.030	0.287
	$\mu_{z_{23}}$	3.883	0.097	0.524		$\sigma_{z_{23}}$	0.942	0.109	0.370
Other non-Aqli	$\mu_{z_{24}}$	4.420	0.093	0.483		$\sigma_{z_{24}}$	1.528	0.031	0.268
Category 3 (dinner-size pizzas)									
Primary Aqli	$\mu_{z_{25}}$	6.163	0.482	1.139		$\sigma_{z_{25}}$	1.437	0.082	0.503
Primary Aqli	$\mu_{z_{26}}$	6.327	0.317	1.022		$\sigma_{z_{26}}$	1.838	0.281	0.684
Primary Aqli	$\mu_{z_{27}}$	5.521	0.431	1.182		$\sigma_{z_{27}}$	0.997	0.196	0.566
Other Aqli	$\mu_{z_{28}}$	5.391	0.506	2.165		$\sigma_{z_{28}}$	0.966	0.406	0.750
	$\mu_{z_{29}}$	5.309	0.478	0.974		$\sigma_{z_{29}}$	1.284	0.298	0.646
	$\mu_{z_{30}}$	2.023	0.302	0.858		$\sigma_{z_{30}}$	0.416	0.180	0.371
	$\mu_{z_{31}}$	4.000	0.497	1.230		$\sigma_{z_{31}}$	0.716	0.253	0.567
	$\mu_{z_{32}}$	5.705	0.404	1.121		$\sigma_{z_{32}}$	0.955	0.224	0.651
Other non-Aqli	$\mu_{z_{33}}$	3.498	0.365	0.853		$\sigma_{z_{33}}$	0.615	0.108	0.392

Table 4

(A) Price sensitivity ($\bar{\alpha}_n$)

	Total	Non-Returners (30)	Returners (70) (Early-returners (33))
Mean	-0.024	-0.027	-0.023 (-0.022)
Standard deviation	0.0074	0.0063	0.0075 (0.0075)

(B) Weight sensitivity ($\bar{\beta}_n$)

	Total	Non-Returners (30)	Returners (70) (Early-returners (33))
Mean	0.0022	0.0014	0.0025 (0.0026)
Standard deviation	0.0030	0.0031	0.0029 (0.0032)

(C) Social-learning intensity (κ_n)

	Total	Non-Returners (30)	Returners (70) (Early-returners (33))
Mean	506.0	471.0	521.1 (529.6)
Standard deviation	45.1	23.1	43.9 (35.6)

(D) Health-damage sensitivity (ω_n)

	Total	Non-Returners (30)	Returners (70) (Early-returners (33))
Mean	-300.9	-302.3	-300.3 (-299.7)
Standard deviation	3.6	3.0	3.7 (3.6)

(E) Risk sensitivity (r_n)

	Total	Non-Returners (30)	Returners (70) (Early-returners (33))
Mean	-714.0	-719.7	-711.6 (-710.9)
Standard deviation	12.0	10.5	11.8 (11.0)

(F) Discounting rate (δ_n)

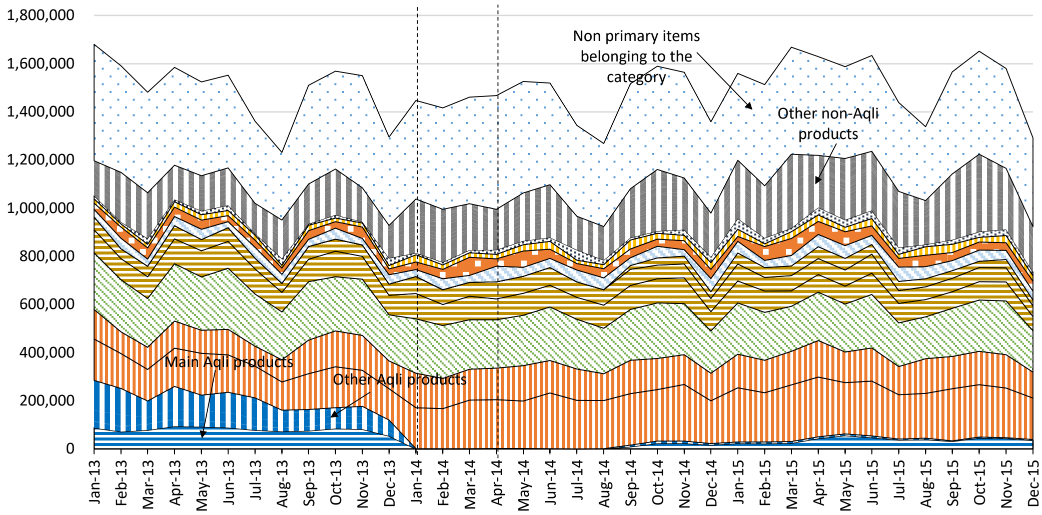
	Total	Non-Returners (30)	Returners (70) (Early-returners (33))
Mean	0.101	0.097	0.102

			(0.104)
Standard deviation	0.005	0.003	0.005
			(0.005)

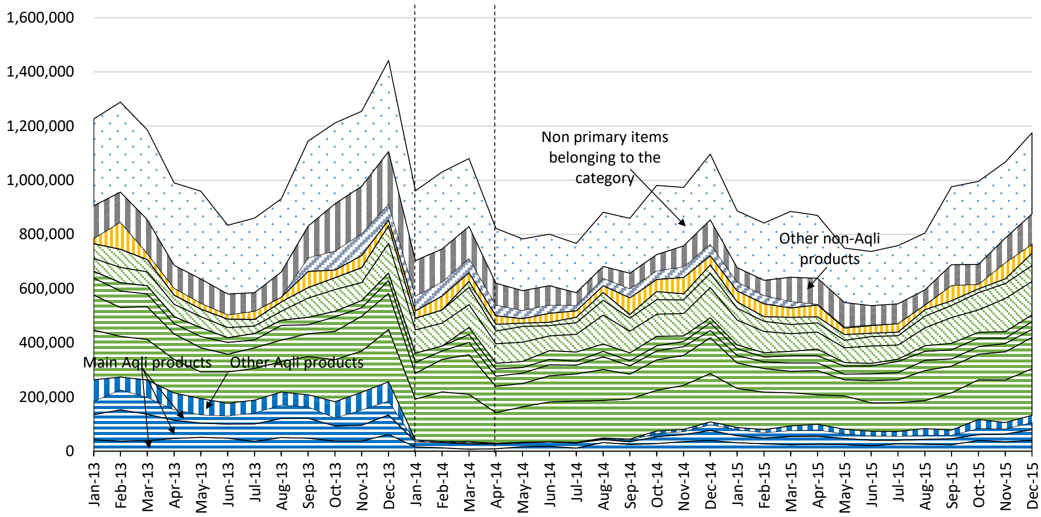
Table 5

	Z_{n1}	Z_{n2}	Z_{n11}	Z_{n12}	Z_{n13}	Z_{n14}	Z_{n25}	Z_{n26}
Non-Returners (30)	6.86	5.99	5.49	5.16	5.77	5.44	6.08	6.26
Returners (70)	7.26	6.08	5.53	5.26	5.80	5.48	6.20	6.35
(Early-returners (33))	7.27	5.48	5.61	5.24	5.74	5.56	6.18	6.30

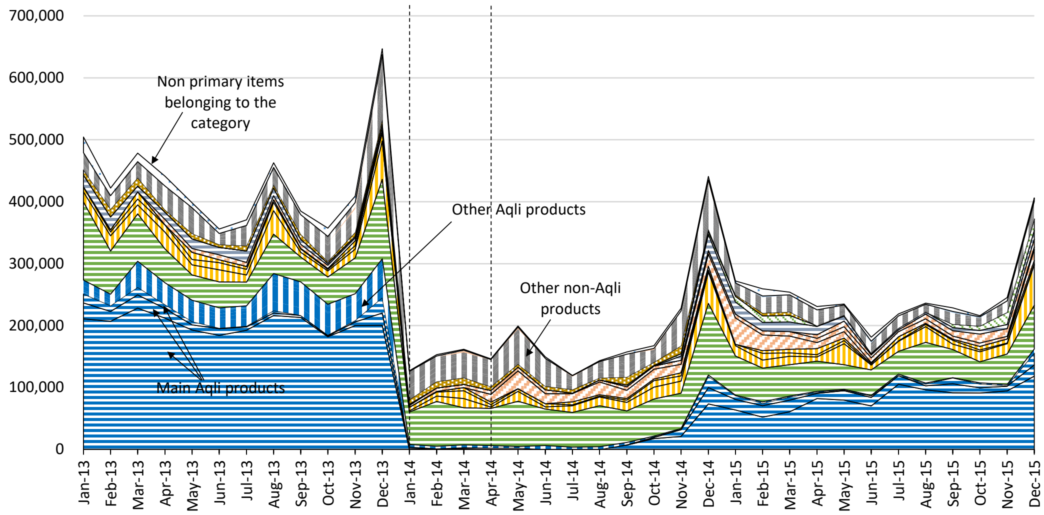
	Z_{n27}	Z_{n28}
Non-Returners (30)	5.52	5.39
Purchasers (70)	5.52	5.39
(Early-returners (33))	5.53	5.39



Category 1 (croquettes)

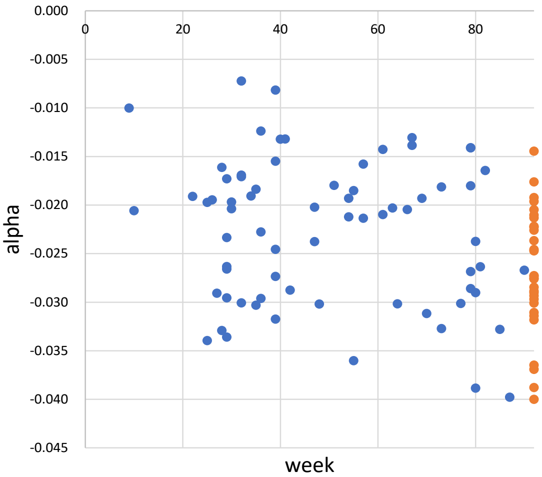


Category 2 (gratins/dorias/lasagnas)

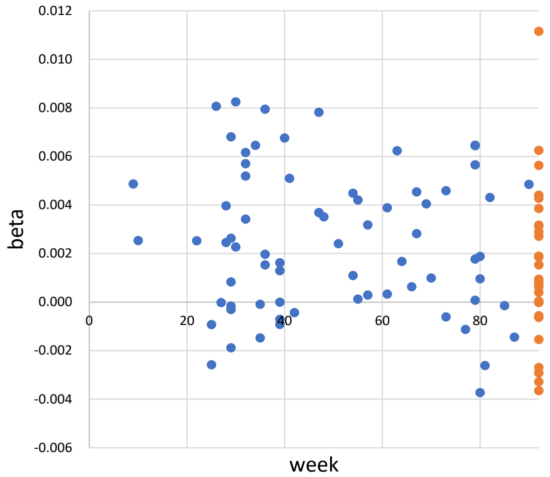


Category 3 (dinner-size pizzas)

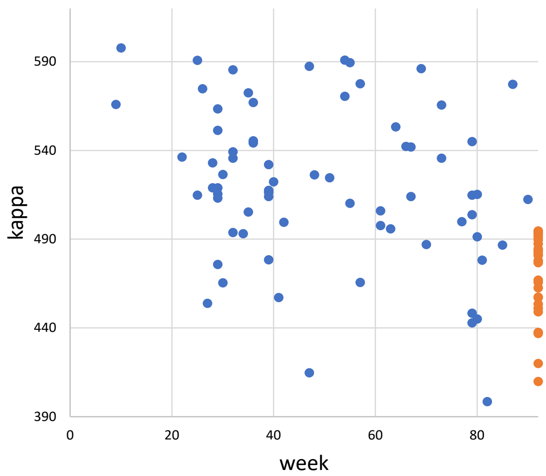
Figure 1



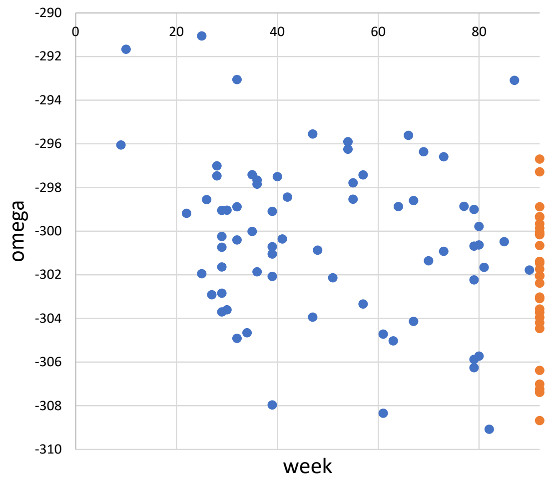
(A) Price sensitivity ($\bar{\alpha}_n$)



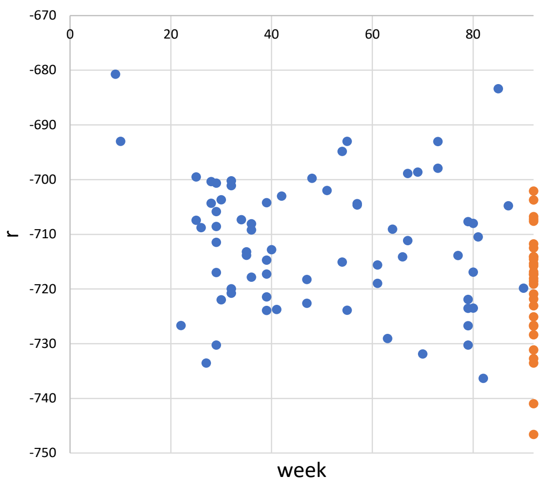
(B) Weight sensitivity



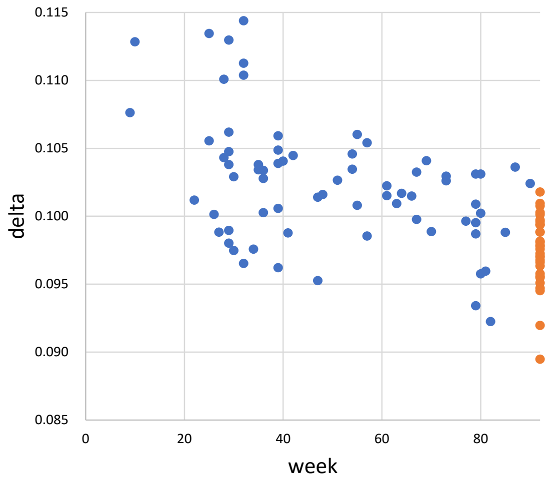
(C) Social-learning intensity



(D) Health-damage sensitivity (ω_n)

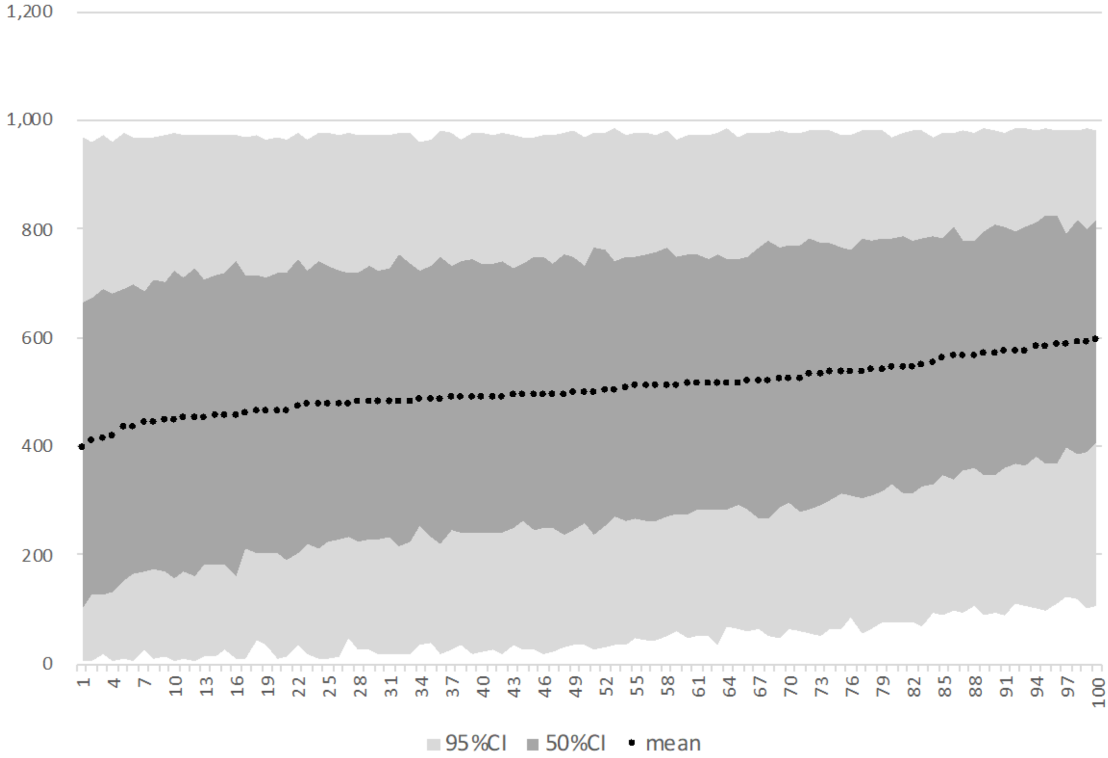


(E) Risk sensitivity (r_n)

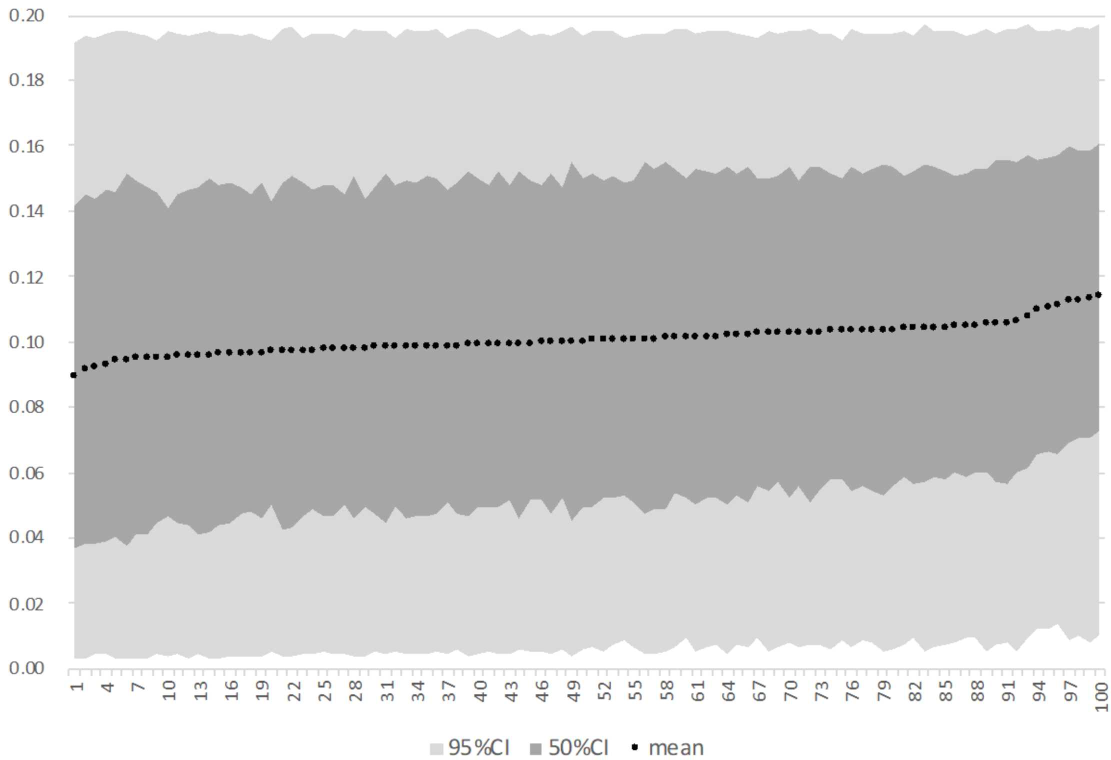


(F) Discounting rate (δ_n)

Figure 2



(A) Social-learning intensity (κ_n)



(B) Discounting rate (δ_n)

Figure 3

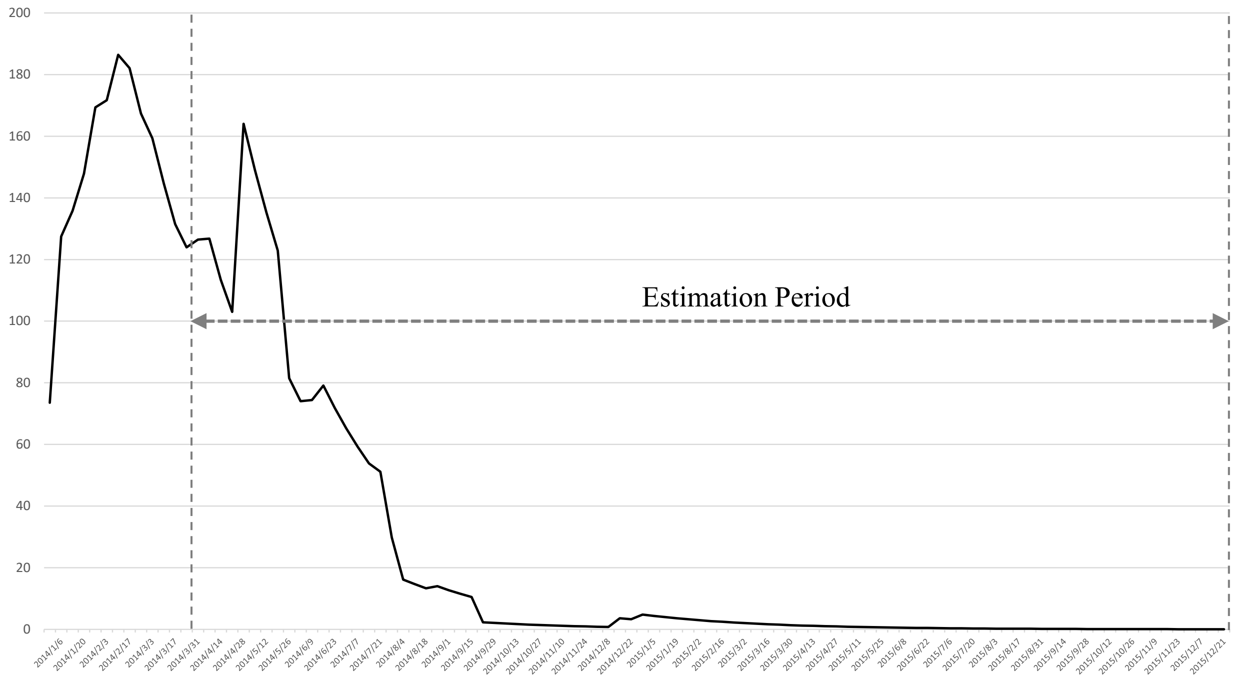


Figure 4