Reputation of Quality in International Trade: Evidence from Consumer Product Recalls

Jiatong Zhong*

This version: October 2018

Click Here for the Most Recent Version

Abstract

This paper quantitatively examines the impact of exporting countries' reputations for product quality on aggregate trade flows. I introduce a novel data set in which recall incidences retrieved from the Consumer Product Safety Commission are matched to U.S. import data from 1990-2009. Using a model of learning I construct a measure for exporter reputation where consumers internalize product recalls as bad signals. Structural estimation of the model finds that reputation is important and especially impactful for products used by children. The market share elasticity of exporter's reputation is around 1.49 across products, similar in magnitude to the average price elasticity, which is around 1.51. Improving reputation can increase export value, but reputation is sluggish: increasing reputation by 10% can take decades for most exporters. Counterfactual exercises confirm that quality inspection institutions are welfare improving, and quality inspection is especially important for consumers of toys.

Keywords: International trade, reputation, Bayesian learning, quality uncertainty, product recalls

JEL Classification: F13,F14,D12,D18,D8,L15

*Department of Economics, Purdue University. zhong60@purdue.edu

1 Introduction

Vertical product differentiation plays a critical role in explaining production and consumption patterns in international trade. The most popular quality measure in trade is price-adjusted sales, which is estimated assuming consumers have perfect information about product quality. Our experiences often depart from this assumption. While a few trade papers have theorized how quality uncertainty affects trade and consumer welfare, their models focus on the static equilibrium outcome (e.g. Bond 1984; Falvey 1989; Chisik 2002) and the empirical investigations are limited to changes after one event, such as implementing quality standards(e.g. Potoski, Prakash 2009). A dynamic model allows demand to be path-dependent and to adjust slowly to quality signals. These properties are important elements in decisions concerning investment into product quality. This paper focuses on the dynamic demand responses and evaluates whether the premises for dynamic quality investment models (i.e. that seller reputation matters) have empirical support in international trade.

When consumers are unsure about quality, they rely on their knowledge of the product, which is referred to as the reputation of sellers. Capturing reputation empirically is challenging for two reasons. First, reputation is history-dependent, so we need to measure a dynamic framework. Second, estimation of a dynamic models needs a data set containing events that repeatedly impact product quality and the market responses to such impacts. This paper proposes a measure of reputation for exporting countries constructed by exploiting the cross-countries, cross-time variation in product recalls. By quantifying the value of reputation, I evaluate exporter incentives to improve product quality. Counterfactual exercises quantify the consumers welfare gains from having an effective quality inspection institution.

This paper introduces a unique data set that merges product recalls with import flows to reveal how market responses to informative signals. I scrape all recall notifications post by the Consumer Product Safety Commission (henceforth CPSC) from 1973 to 2015, and use recall date, product descriptions, and exporting countries to match the recall incidences to U.S. monthly import data from April 1990 to December 2009.[1] A prominent data pattern revealed in this data set, as illustrated by figure A.1, is that larger exporters tend to face more recalls. However, if we zoom in and focus on one exporting country, its market share declines immediately after a major recall event hits.[2] An intuitive explanation is that volume matters: conditional on the fraction of unsafe products, countries selling more units are more likely to face recalls, so even if recalls have negative impacts on sales, the effect is obscured by sales

[1]The Commission provides public access to their recall database through a Recalls Application Program Interface (API).

[2]See figure A.2 for an example using Hong Kong export of toys.

itself in a micro-econometric analysis. This paper disentangles the impact of recalls from the sales volume and provides a quantitative method to evaluate the impact of bad signals.

In this model, each product-exporting country pair is a variety, and each variety has a different fraction of safe products. Consumers do not know the fraction for each variety, and unsafe products look identical to safe products before purchase. However, they can use observed recalls to learn about the fraction and form an expectation for the quality which enters aggregate demand as a product characteristic. Following Board and Meyer-ter-Vehn (2013), I define reputation as the expected quality formed in each period of the learning process.

The model is estimated exploiting the market share responses to recalls, and the mean and variation of recalls. The parameters that shapes consumer learning process are identified with a convergence property of Bayesian learning. Learning parameters are estimated such that the mean and variance of recalls predicted by the learning outcomes of the last period match the moments from the observed recalls. Reputation is constructed with the learning parameters, quantity of imports, and recalls. The taste for reputation is estimated such that predicted market shares match observed market shares. All parameters are estimated simultaneously using generalized method of moments, and as a mathematics program with equilibrium constraints (MPEC).

Using estimated reputation and preferences, I perform counterfactual exercises that concern exporters and consumers respectively. First, I calculate the impact of recall events on market share and trade value. Estimates suggest that consumers do not factor reputation into decision making for some products (like lamps), but they weight reputation for products like toys and children's clothes heavily. On average, a 10% improvement in reputation can increase market share by 14.9%, and for some products the increase can be up to 50.7%. However, reputation is sluggish, especially for small exporters who used to be large exporters. Even for an average large exporter, it takes almost 36 years of recall-free presence in the United States to improve its reputation by 10%.[3] Second, I examine the value of having a quality inspection institution by simulating a counterfactual scenario in which the probability of a bad product being recalled is reduced from 90% to 50%. Average welfare losses vary from 0.028 cent to 87 cents per purchase depending on the type of product. Total welfare losses can average up to 2.78 billion dollars a year for toys, while for all other products (e.g. sweaters and battery) the mean is typically less than a million dollars a year.[4] The results suggest that for the importing country, a product quality inspection institution like the CPSC can improve consumer welfare.

This paper is related to the trade literature involving quality uncertainty. The theoretical component uses a learning approach, which adds to two popular methods to model quality

[3]Here, large is defined as in the upper quartile of export quantity.

[4]The United States spent over 20 billion dollars on toys last year.

uncertainty of imported goods: adverse selection (Bond 1984; Donnenfeld et al. 1985; Donnenfeld 1986a; Donnenfeld, Mayer 1987) and reputation premium (Falvey 1989). I introduce a dynamic framework featuring quality uncertainty into the international context, which is closer to the recent models of reputation and uncertain product quality developed in industrial organization literature. The learning model also allows me to evaluate the welfare impact of information disclosure, which only a few papers have theorized (Creane 1998; Creane, Miyagiwa 2008), and I am not aware of any paper that quantifies it.

This paper contributes to the relatively small empirical literature of quality uncertainty in trade, which primarily examines the effects of national and international quality standards (Swann et al. 1996; Potoski, Prakash 2009). The empirical component of this paper departs from that approach in two ways. First, in my model the signals are sent repeatedly, so I propose an empirical strategy that can handle a dynamic framework. Second, I introduce a new source of data that reflects product quality. Compared to quality standards, recalls provide more frequent changes to infer reputation. Relative to the customer ratings from online platforms used in empirical industrial organization (for example Mayzlin et al. 2014), this data set contains more products and information about exporting countries.[5]

The empirical analysis also contributes to studies using product recall data. Freedman et al. (2012) used toys recall data from the CPSC to run a difference-in-difference regression, estimating the spillover effect in volume of sales to the producer and the industry. Grundke, Moser (2014) examines whether FDA uses import refusals strategically during recessions under the pressure of protectionism. This paper offers a new topic and a corresponding empirical method to utilize information from recall data.

The model builds on a rich literature studying sellers' reputation when product quality cannot be perfectly observed (See Bar-Isaac et al. (2008) for a detailed survey). It fits into the branch of literature where sellers have hidden information from consumers, and it is most similar to that in Bar-Isaac (2003), sharing the feature of learners updating their belief under Bayes rule. It borrows the definition of reputation from Board and Meyer-ter-Vehn(2013) because they explicitly model signals in a manner close to how product recalls happen. This paper focuses on consumer responses instead of firms' investment in product quality, which is a common interest in the learning literature in industrial organization. Another important difference is how information is distributed: I abstract away from the concept "experimentation" discussed in Rothschild (1974) and Bolton, Harris (1999), which features consumers strategically making purchase decisions in order to obtain more information. In my context, signals are sent out by a quality inspection institution. The empirical literature on sellers' reputation

[5]See Donnenfeld (1986b); Falvey, Kierzkowski (1984) for additional empirical works on quality standards. Chen et al. (2016) uses online reviews as a proxy for reputation of foreign individual sellers.

uses almost exclusively data from electronic market place. My results are consistent with their conclusions, that sellers are rewarded by having a good reputation (e.g. Eaton 2005), although it is not the case for all products.[6] Most empirical works cover one specific good or service (e.g. iPod in Saeedi (2014)), but my study covers many products, and studies impact on exporters instead of individual sellers.

This paper contributes to the growing research applying learning models in trade, which mostly concerns how firms learn about foreign markets before entry (Eaton et al. 2009; Albornoz et al. 2012; Holloway 2017) and how firms building a relationship with foreign suppliers (Rauch, Watson 2003). Two learning models are popular among trade economists, learning with experimentation featuring firms start with small transactions before expansion (Albornoz et al. 2012; Rauch, Watson 2003) and Bayesian learning characterizing how firms obtain information about foreign markets (Eaton et al. 2009; Holloway 2017). This paper follows the tradition of Eaton et al. (2009) and Holloway (2017), but focuses on the consumers' perception.

This paper is organized as follows. In the next section I introduce a partial equilibrium model that captures how consumers update their perception of an exporter's reputation in a market using observations of product recalls in a period. Section 3 explains the empirical strategy for estimating this model. Section 4 describes the novel data set, and I report the results in section 5 and 6. Section 7 concludes.

2 A Learning Model for Exporters' Reputations

In this model, I introduce the definition of reputation, how it evolves over time, and how the market responds to it, focusing on the consumers' decisions. I assume that firms within an exporting country face perfect competition, and supply inelastically in each period. Consumers make purchase decisions based on prices and the current reputation for each exporting country. After purchase, they observe quantity sold, recalls, and update the reputation at the end of the period with past reputation and the new signals they observe.

2.1 Consumers' Problem

There is a continuum of consumers indexed by i. In each period t, each consumer consumes one unit of a differentiated product, s, and $y_{i,t}$ units of a homogeneous product. Consumers do not observe the true quality of the differentiated product, but they observe the country-of-origin, j. The differentiated product is either safe or unsafe, characterized by the unobserved quality z that takes value 1 if it is safe, and 0 otherwise. Consumers cannot distinguish between

[6] Other papers that have similar conclusions include Livingston (2005); Houser, Wooders (2006); Mayzlin et al. (2014); Wu (2016); Saeedi (2014)

safe and unsafe products before purchase, but they observe the outcome after purchase which factors into their realized utility. I take an utility function similar to that in Petrin (2002). The utility after purchase and quality revelation is written as

$$u_{ijs,t} = \alpha_0^s \log(y_{i,t}) + \alpha_x^s z_{js,t} + \eta_{js} + \psi_{s,t} + \xi_{js,t} + \epsilon_{ijs,t}.$$

η_{js} is the time-invariant preference common across all consumers for a product from a country, which captures time-invariant unobserved characteristics, such as Italian men's wool suits are considered better. ψ_{st} captures the time specific demand for product s, for example higher demand for toys in the last quarter of the year. $\xi_{js,t}$ represents unobserved demand shocks like retail channels and unobserved variety characteristics. $\epsilon_{ijs,t}$ is the idiosyncratic preference shock that follows i.i.d. Extreme Value distribution.

In each period, consumers maximize their expected utility by choosing one exporting country to buy one unit of differentiated product from. Let $\mathcal{H}_t$ denote the information set available to consumers when making a purchase decision. The expected quality of product s from country j is denoted as

$$x_{js,t} = \mathbb{E}[z_{js}|\mathcal{H}_t].$$

We will discuss what is in the information set $\mathcal{H}_t$ and the functional form of expectation in the next section. Using the law of iterated expectations, we can write consumer's maximization problem as:

$$\begin{aligned} \max_{j \in \mathbb{J}_s} \quad \mathbb{E}[u_{ijs,t}] &= \mathbb{E}\left[\mathbb{E}[u_{ijs,t}|\mathcal{H}_t]\right] \\ &= \alpha_0^s \log(y_{i,t}) + \alpha_x^s x_{js,t} + \eta_{js} + \psi_{st} + \xi_{js,t} + \epsilon_{ijs,t} \qquad (1) \\ \text{s.t.} \quad & y_{i,t} + p_{js,t} \leq I_t, \end{aligned}$$

where I_t is the budget constraint that can be interpreted as income, $p_{js,t}$ is the price for one unit of differentiated product s from country $j \in \mathbb{J}_s$, and $\mathbb{J}_s$ is the set of exporters who sell product s to the United States. Price of the homogenous product is normalized to 1. The consumer optimization problem is a standard discrete choice problem as in Petrin (2002), where expected quality of the differentiated product enters consumer's decision as a product characteristic. Following Board, Vehn Meyer-ter (2013), I refer to the expected quality $x_{js,t}$ as the reputation for product s from country j at period t, and I will henceforth call it "reputation".[7] In the next section, I will derive the law of motion of reputation.

[7] Note that the definition of reputation is similar to that in the "perfect bad signal" scenario in Board, Vehn Meyer-ter (2013), but the model is different in two ways. First, this model is in discrete time while Board, Vehn Meyer-ter (2013) sets their model in continuous time. More importantly, Board, Vehn Meyer-ter (2013) concerns firm's investment in efforts and their model includes a productivity shock, but this model abstracts away from firm's strategy or productivity.

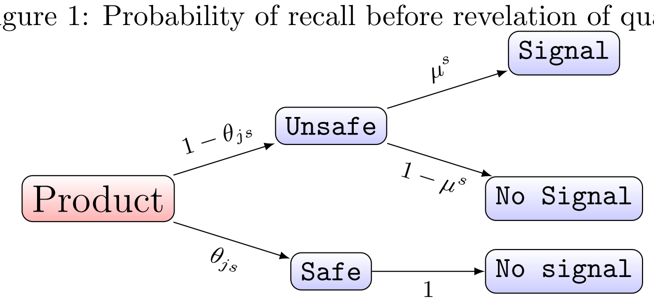
2.1.1 Deriving the updating process

This section begins with a sketch of the probability problem a consumer faces when she infers the expected quality of the product using history of sales, recalls, and country-of-origin. I then derive the reputation updating process from the consumer's rational expectation, and show that reputation can approach the true average quality for each exporting country given sufficient periods of learning.

2.1.2 Deriving the updating process

Consumers do not observe quality of differentiated products, but they can observe the country-of-origin label. I assume that the fraction of safe product s from an exporting country j is θ_{js}, which consumers do not know fully, but they can learn about it through signals. In particular, their belief follows a distribution on [0,1], and signals change that distribution over time. The true fraction is assumed to be constant over the periods of learning. If the product is unsafe, then there is a probability μ^s that it will be recalled. That probability is product-specific, but common across time and across exporting countries. Figure 1 illustrates the above-described process.

Figure 1: Probability of recall before revelation of quality

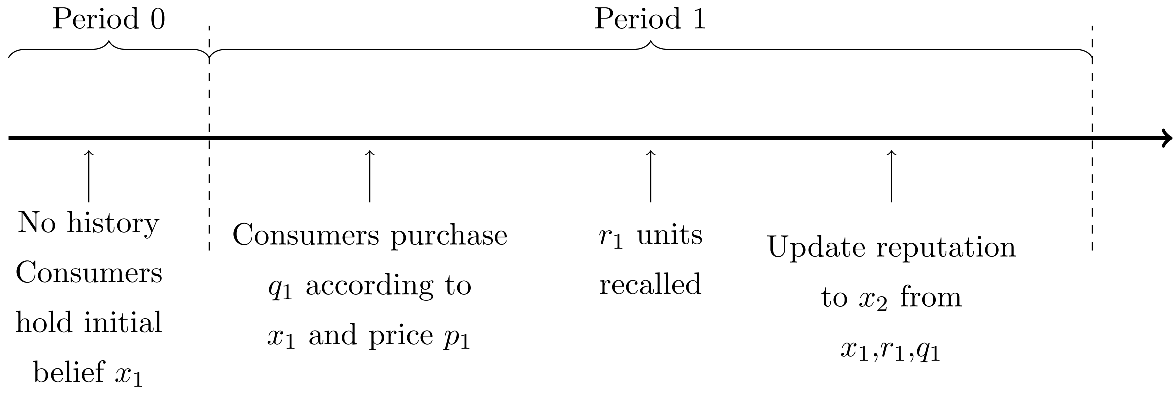


I assume that safe products will never result in a recall, which should not be far from truth. Most recalls are triggered *after* one or more hazardous events are reported by consumers or retailers. CPSC investigates the reports and if the Commission decides that there is a "substantial product hazard", it will issue a recall. If a retailer or manufacturer voluntarily recalls the product—usually after a consumer complaint—the recall notice will be issued faster.[8] In both cases, recalls are mostly complaint-driven, so it is reasonable to say that recalls are only issued to problematic products.

Consumers do not know the value of θ_{js}, but they form an expectation of its value based on informative signals. Their information set for product s from exporter j at period t is $\mathcal{H}_{jst}$,

[8] Consumers, government agencies and medical practitioners can voluntarily file reports of product hazards to the CPSC, while manufacturers, importers, distributors, and retailers have a legal obligation to report the products to the CPSC once they learned the product defects and hazards.

Figure 2: Timing of the consumer reputation update process across periods



which contains the history of recalls $\{r_{js,\tau}\}_{\tau=1}^{t-1}$, quantity $\{q_{js,\tau}\}_{\tau=1}^{t-1}$, and reputation $\{x_{js,\tau}\}_{\tau=1}^{t-1}$ at period t. When realized quality is 0 or 1, the reputation coincides with the expected fraction of true products:[9]

$$x_{js,t} = \mathbb{E}[z_{js}|\mathcal{H}_t] = \mathbb{E}\left[\theta_{js}|\mathcal{H}_{jst}\right].$$

Consumers' expectations form the menu of reputation $\{x_{js,t}\}_{j\in\mathbb{J}^s}$ for different exporters j. Information set $\mathcal{H}_t$ contains all information sets for a particular country-product pair $\mathcal{H}_{jst}$, but the information necessary to update one country's reputation is only its own history.

Figure 2 illustrates the timing of events in the first two periods, and other periods follow the same pattern. Before they make the purchase decision in period t, consumers learn about the probability of getting a safe product if they buy from country j by Bayesian updating their probability assessment using the signals of recalls they receive in last period.

Purchasing from country j is analogous to making a random draw from a pool of size $q_{js,t}$. Given that the true and unobserved fraction of safe products is θ_{js} for a country j, consumers purchased a total of $q_{js,t}(1-\theta_{js})$ units of unsafe products. For each unit of unsafe product, there is a probability μ that the CPSC will issue a recall. This can be due to consumers being unaware of the product defect or the CPSC's investigation failing to confirm the product's defect after the initial report.

In derivation of updating process, I suppress product and country indices since the same process applies to all product-country pairs. The signals are sent through standard Bernoulli trials, and, following Bayes' rule, the likelihood function with a data realization is:

$$\rho(r|\theta) = \mathcal{L}(\theta) \propto [(1-\theta)\mu]^r[1-(1-\theta)\mu]^{q-r} = \gamma^r(1-\gamma)^{q-r}$$

[9] More generally, when realized quality is a when product is unsafe and b when product is safe, $b > a$, expected quality is a linear transformation of the conditional expectation of θ_{js}: $x_{js,t} = \mathbb{E}[z_{js}|\mathcal{H}_t] = a+(b-a)\,\mathbb{E}\left[\theta_{js}|\mathcal{H}_{jst}\right]$. The reputation motion is a straightforward extension of the current form.

where $\gamma \equiv (1-\theta)\mu$ is defined for notation simplicity. γ is the unconditional probability of sending a recall signals for each draw.

If we assume that the prior distribution of γ is a Beta distribution, the reputation updating process follows the equations in Proposition 1. The Beta distribution is a conjugate prior distribution for the Bernoulli likelihood function: it means that before and after the update, the distributions of γ are both Beta distributions. This is algebraically convenient for us to compute an expectation before and after learning in a period. [10]

Proposition 1. *When we choose a Beta distribution $\mathcal{B}(\beta_0, \delta_0)$ as the prior distribution for $\gamma \equiv (1-\theta)\mu$, the reputation update from period t to $t+1$ follows:*

$$\begin{cases} x_1 = 1 - \dfrac{\beta_0}{\mu(\beta_0+\delta_0)} & \text{(2a)} \\[2ex] x_{t+1} = \dfrac{\beta_t+\delta_t}{\beta_t+\delta_t+q_t} x_t + \dfrac{q_t}{\beta_t+\delta_t+q_t}\left(1 - \dfrac{r_t}{\mu q_t}\right) & \text{(2b)} \end{cases}$$

with β_0 and δ_0 as the initial parameter values for the Beta distribution $\beta_t = \beta_0 + \sum_{\tau=1}^{t-1} r_\tau$ and $\delta_t = \delta_0 + \sum_{\tau=1}^{t-1} q_\tau - \sum_{\tau=1}^{t-1} r_\tau$.

The intuition for β_0 is the cumulative units of goods ever recalled from a variety before the first period the data set allows econometricians to observe. Similarly, δ_0 is the cumulative units of un-recalled products sold into the United States before the first observation. β_0 and δ_0 absorb the history before the starting period in estimation. β_t is the total cumulative units of recalled products up to period t, and δ_t is the total cumulative units of safe products sold up to t. The summation of δ_t and β_t produces the total cumulative units of goods sold before period t.

Equation 2b is in the form of a weighted average of current reputation x_t and new information $1 - \dfrac{r_t}{\mu q_t}$. The first term in equation 2b contains a coefficient of x_t that captures the persistence of reputation. The coefficient can be re-written in the form:

$$\frac{\beta_t+\delta_t}{\beta_t+\delta_t+q_t} = \frac{\beta_0+\delta_0+\sum_{\tau=1}^{t-1} q_\tau}{\beta_0+\delta_0+\sum_{\tau=1}^{t} q_\tau}.$$

The denominator of coefficient is the cumulative units of goods sold at the end of period t, and numerator is the cumulative units sold before period t, so intuitively, the coefficient captures the "weight of history". When β_0 and δ_0 are small relative to the total quantity sold in past periods, the coefficient is dominated by the fraction of the summation of the units sold up to

[10] In appendix C.2 I include a discussion of using truncated Beta distribution as a prior, for readers who are concerned about the upper limit of the distribution of γ. I concluded that if β and δ are large enough, the reputation updating procedure is the same as the one shown using standard Beta as a prior.

period $t-1$ over the units sold up to period t. This weight is between 0 and 1, and it increases over time, so it is a term that captures the convergence of reputation.

The second term captures the new information in period t. The coefficient is the fraction of quantity sold in period t in the cumulative units of goods sold at the end of period t, which is intuitively the "weight of new information". The term $\left(1 - \frac{r_t}{\mu q_t}\right)$ is the expected fraction of safe products in the market in period t.

Equation 2a represents the initial condition. $\frac{\beta_0}{\beta_0 + \delta_0}$ is the fraction of the cumulative sum of recalled products relative to the sum of all units sold before the first observation. Adjusted by the efficiency of the recall $\frac{1}{\mu}$, $\frac{\beta_0}{\mu(\beta_0 + \delta_0)}$ is the expected fraction of unsafe products in the first period.

β_0 and δ_0 must be positive numbers, as implied by intuition, and they are likely in a magnitude comparable to (or larger than) the volume of trade flows observed. The probability of recall (given that a unit of product is bad) is given by parameter μ, and $\mu \in (0, 1]$. μ cannot be zero; otherwise, equation 2a and 2b are not well-defined. Intuitively, the effectiveness of inspection cannot be so bad that a recall is close to impossible. Quantity q_t is a positive number that does not go to infinity, and the units of recall r_t are nonnegative and bounded above by q_t in each period. The range of parameters in proposition 1 imposes almost no other restrictions beyond those implied by economic intuition, but they are necessary for the asymptotic property presented in the next section.

2.1.3 Asymptotic property of reputation learning

Bayesian learning is a type of perfectly rational learning. With some restrictions, the expectation converges asymptotically to the true value agents learn about. I will refer to this asymptotic property as "effective learning" henceforth. I will return to this property in the estimation section, as it is useful for identification.

I assume that, conditional on the history $\mathcal{H}_{jst}$, the fraction of safe products θ_{js} and probability of recall for unsafe products μ^s, the expectation of import in period $t+1$ is product-country-specific, but time-invariant. That is, consumers do not learn about the size of market from history. This assumption and the assumption on bounds of parameters are formalized in Appendix C.6 as assumption 1 and 2. Together, they provide sufficient conditions for asymptotic effective learning.

Theorem 1. *Given assumptions 1 and 2, learning is effective asymptotically. That is, the expectation converges to the truth when T is large:*

$$x_{js,T} \to \theta_{js}, \quad \textit{as} \quad T \to \infty$$

Proof. See Appendix C.6.[11] □

In each period t, every consumer forms their expectation for product quality from the observed signals $r_{js,t}$ and market size $q_{js,t}$, and then from the menu of reputation and price they make their purchase decision. By aggregating individual purchase decisions, we can compute the countries' market shares using a discrete choice model.

2.2 Equilibrium

Following standard logistic demand assumptions and let the budget constraints hold with equality, the market share of country j in a particular product market s in time t is:

$$\begin{aligned} s_{js,t} &= \int_{\epsilon_{ijs,t} | u_{ijs,t} > u_{ij's,t} \forall j' \neq j\}} d\mathcal{F}_\epsilon(\epsilon) \\ &= \frac{(I_t - p_{js,t})^{\alpha_0^s} \exp(\alpha_x^s x_{js,t} + \eta_{js} + \psi_{st} + \xi_{js,t})}{1 + \sum_{j'=1}^{J_s} (I_t - p_{j's,t})^{\alpha_0^s} \exp(\alpha_x^s x_{j's,t} + \eta_{js} + \psi_{st} + \xi_{js,t})} \end{aligned} \tag{3}$$

subject to constraint:

$$I_t \geq p_{js,t}$$

In equilibrium, the goods market clears. In each period, the United States imports as many units of products from each exporter as demanded in the domestic market. The United States is treated as a supplier as well, and the utility of purchasing from the U.S. is normalized to 1. Since firms are perfectly competitive within an exporting nation, price is determined by country-specific costs and treated as given in this framework.

Formally, the equilibrium definition is:

Definition 1 (Equilibrium with Learning). *An equilibrium in this model is defined as a $J \times S \times T$-by-3 matrix of price, reputation and import flows $[p_{js,t}, x_{js,t}, q_{js,t}]$ with a Bayesian learning motion such that:*

1. *Import Market Clears:*

$$S_{js,t} = s_{js,t}(p_{js,t}, x_{js,t}, \xi_{js,t}; \alpha^s, \mu^s, \beta_0^s, \delta_0^s)$$

2. *The Bayesian learning motion satisfies:*

$$x_{js,t+1} = \frac{\beta_{js,t} + \delta_{js,t}}{\beta_{js,t} + \delta_{js,t} + q_{js,t}} x_{js,t} + \frac{q_{js,t}}{\beta_{js,t} + \delta_{js,t} + q_{js,t}} \left(1 - \frac{r_{js,t}}{\mu_{js,t}^s}\right)$$

where $\beta_{jst} = \beta_0^s + \sum_{\tau=1}^{t-1} r_{js,\tau}$ and $\delta_{jst} = \delta_0^s + \sum_{\tau=1}^{t-1} q_{js,\tau} - \sum_{\tau=1}^{t-1} r_{js,\tau}$; and β_0^s and δ_0^s as the initial parameter values.

[11] This proof is only slightly different from a standard proof of convergence in Bayesian learning.

3 Empirical Strategy

Income I_t, price $p_{js,t}$, total units of sale $q_{js,t}$, quantity of risky products $r_{js,t}$, and market share $S_{js,t}$ are data in the equilibrium with learning. For each product s, μ^s, β_0^s, δ_0^s, and the vector of demand function coefficients α^s are parameters that need to be estimated. Price, quantity, the number of recalls, and market share vary across time and varieties, while income varies over time only. Parameters vary across products, but are constant over time and across exporters.

The baseline estimation is done product by product. A product is a commodity classified under a six-digits harmonized system code in the import data. Within each product s, the set of learning parameters $(\mu^s, \beta_0^s, \delta_0^s)$ enters the model non-linearly, and given estimated reputation $\{x_{js,t}\}_{j\in J^s, t\in 1,2...T}$, the vector of demand parameters α^s enters linearly.

There are three main challenges to estimation. First, although the demand equation can be linearized, the system of equations is still non-linear because of the Bayesian learning motion. In addition, the reputation measure $x_{js,t}$ is constructed, so to make sure its value aligns with the data, I use a property of Bayesian learning and introduce an additional objective function in estimation. Multiple objective function optimization problem (henceforth MOOP) is common in engineering, but less so in economics, so I borrow a classic method in engineering to transform this problem into a single objective function problem. Finally, price is endogenous in the demand equation, so it calls for an instrumental variable.

The empirical strategy has two parts, though they are estimated simultaneously. These parts correspond to the main challenges in identification. In the first part, I use history of import quantity and recall units to back out the parameters that determine reputation dynamics, exploiting the asymptotic property in theorem 1. This condition implies that after enough periods of learning, the reputations for each country approach the true unobserved fraction of good products.[12]

3.1 Estimating Bayesian Updating Parameters from Recalls and Quantities

Separately identifying preference for reputation α_x and the probability of a recall μ requires us to take advantage of a property of learning, because reputation $x_{js,t}$ is constructed. Intuitively, I use the fraction of unsafe products implied by the learning model to predict the mean and variance of recalls, and match the moments to those observed in recall data. Theorem 1 shows that, given enough periods of learning, reputation converges to the true expected quality. I take the vector of reputation in the last period $x_{js,T}$ and use it as a proxy for the unobserved fraction of good products θ_{js}. To ensure that consumers actually learn sufficiently,

[12] In the estimation, "enough" is defined as at least 10 quarters of learning.

I only include exporters who have been in the U.S. market for more than 10 quarters. Using J'_s to denote the set of exporters of product s that we have observed for more than 10 periods, we can formulate this criteria as the following likelihood estimation. Given the units of import from each country in each period $q_{js,t}$, the number of unsafe products in the market in period t is:

$$L_{s,t}(\mu^s, \beta_0^s, \delta_0^s) = \sum_{j \in J'_s} q_{js,t} \times x_{js,T}(\mu^s; \beta_0^s, \delta_0^s)$$

I observe the total number of recalled products in each period $R_{s,t} = \sum_{j \in J'_s} r_{js,t}$. For each lemon in the market, the probability of being recalled is μ^s. $R_{s,t}$ is the realization in period t of $L_{s,t}$ independent Bernoulli trials with "success" probability μ^s and follows binomial distribution. Given that $L_{s,t}$ is large, we can use a normal distribution $\mathcal{N}(\mu^s L_{s,t}, \mu^s(1-\mu^s)L_{s,t})$ to approximate the binomial distribution, and the log-likelihood function is:

$$\mathcal{L}(R_{st}|\hat{\theta}(\mu^s; \beta_0^s, \delta_0^s), Q_{s,t}) = \sum_{t=1}^{T} log\phi(R_{st}|\hat{\theta}(\mu^s), Q_{s,t}) \tag{4}$$

where $\phi(R_{st}|\hat{\theta}(\mu^s), Q_{s,t})$ is the normal probability density function with mean μL_t and variance $\mu(1-\mu)L_t$.[13] Given learning parameters, reputation can be constructed without price or market share data.

3.2 Demand Estimation

For each set of value $(\mu^s, \beta_0^s, \delta_0^s)$, reputation can be computed as a given product's characteristics. The rest of the parameters—the preference parameters (α_0^s, α_x^s) constants and fixed effects—are estimated from a standard discrete-choice demand system. I follow Khandelwal (2010) and treat purchasing from the United States as the outside option in the discrete choice. In cases without income heterogeneity, the demand equation can be linearized (see (Berry 1994)). The log-linearization of market share equation 3 is:

$$\ln(s_{sj,t}) - \ln(s_{s,US,t}) = c^s + \alpha_x^s x_{js,t} + \alpha_0^s \ln(I_t - p_{js,t}) + \eta_{js} + \psi_{st} + \xi_{js,t}$$

I_t is the average household expenditure on consumption goods per quarter over all observed periods. The coefficient α_0^s is the own price elasticity of the good s. The term $\ln(I_t - p_{js,t})$, given price is involved, is correlated with the unobserved product characteristics. Thus, I use unit shipping cost as the price instruments following Khandelwal's argument (Khandelwal 2010).[14]

[13] The approximation is mostly for computation. Matlab cannot compute the likelihood of this binomial distribution since the power exponent is too large.

[14] Khandelwal provided an explanation for the validity of these instruments, see (Khandelwal 2010) for details. I have also tried exchange rates and oil price times distance between importer and exporter as instruments, but

The definition of market share as a fraction of trade values instead of quantity implies a small modification of the linearized equation. The regression equation in the case of homogeneous income is:

$$\ln(S_{js,t}) - \ln(S_{US,s,t}) - \ln(p_{js,t}) = c^s + \alpha_x^s x_{js,t} + \alpha_0^s \ln(I_t - p_{js,t}) + \eta_{js} + \psi_{st} + \epsilon_{js,t} \qquad (5)$$

Denote $y_{js,t} \equiv \ln(S_{js,t}) - \ln(S_{US,s,t}) - \ln(p_{js,t})$, and henceforth I will use $y = \{y_{js,t}\}_{s,t}$ to refer the dependent variable constructed from market shares.

The residual of regression 5 forms the orthogonality condition necessary for GMM estimation:

$$\mathbb{E}[\xi_{js,t}|h(x_{js,t}, z_{js,t})] = 0$$

where h is a function of the observed exogenous variables and the instrument.[15] The moment condition for the GMM estimator is:

$$\begin{aligned} g(\hat{\alpha}) &= \frac{1}{T \times K} \sum_{t=1}^{T} \sum_{k=1}^{K} \hat{\xi}_{k,t} \cdot h(z_{k,t}, x_{k,t}) \\ &= \frac{1}{T \times K} \sum_{t=1}^{T} \sum_{k=1}^{K} Z' \hat{\xi}_{k,t} \\ &= \frac{1}{T \times K} \sum_{t=1}^{T} \sum_{k=1}^{K} Z'(y_{k,t} - X_{k,t}\hat{\alpha}) \end{aligned}$$

In the baseline estimation, Z is a simple vector of exogenous variables and instruments.[16]

---

the first stage test shows that they are not as ideal.

[15]In the Nested Fix Point approach (Berry 1994), the unobserved characteristic ξ_t is calculated by inverting the market share equation 3 . The MPEC approach does not require such an inverse and can thus be faster.

[16]In the main estimation, I provided the constraint Jacobian and Hessian matrix to improve computation speed. Given that the variable $x_{js,t}$ is a non-linear function of some parameters, the Jacobian and Hessian matrix will be much more complicated. I also tried using $h(.)$ as a second order polynomial following Dubé, Fox and Su without providing the Jacobian. The estimation results for one industry are similar to that using only simple instruments.

Newey (1990)discusses finding asymptotically efficient instruments for nonlinear models using nonparametric method.He introduced two methods: k-nearest neighborhood and series approximation–which is the polynomial-based instruments. Series approximation is more suitable in this case because I provide constraint Jacobian to speed up computation. To derive the constraint Jacobian I need the optimal set of instruments to be differentiable. In fact, this set of instruments performs reasonably well in an efficiency comparison. Reynaert and Verboven (Reynaert, Verboven 2014) ran a simulation estimating a random coefficient model and found that the set of instruments used in Dubé, Fox and Su outperforms pseudo Monte Carlo integration.

3.3 Estimating the Model as One MPEC Problem

The model is estimated as a mathematical program with equilibrium constraints (henceforth, "MPEC") problem. This is a technique widely used in engineering and recently adopted in industrial organization to solve optimization problems with many nonlinear constraints.[17] Dubé, Fox and Su have shown that MPEC has a significant speed advantage for the estimation of large-dimensional problems with many markets (Dubé et al. 2012) and also improves convergence compared to the nested-fixed point algorithm. By setting the Bayesian learning procedures as dynamic constraints, the model can be estimated simultaneously as a MPEC problem.

This problem is also a Multiple-Objective Optimization Problem as we have both the GMM objective function and the maximum likelihood function introduced in section 3.1. The MLE adds a layer of complication to the econometrician's problem, but is necessary to pin down the structural parameter μ. I used the epsilon-constraint method for MOOP first introduced by Haimes (1971) to re-write the MLE objective function as an inequality constraint. The epsilon-constraint method keeps one of the objective functions and rewrites the rest into constraints by restricting them within an econometrician-specified range from their optimal values. Before the estimation, the econometrician must run the optimization problem as a single-objective function problem to obtained the objective values for each objective function, which is the "optimal value" mentioned above. Intuitively, there is a trade-off in optimization when there are multiple objective functions. The epsilon-method prioritizes one objective function as long as the secondary objectives are "good enough."[18] The inequality constraint introduced by this method is:

$$|\mathcal{L}(R_t|\hat{\theta}(\mu;\beta_0,\delta_0),Q_t)-\mathcal{L}^*|\leq\epsilon$$

in which $\mathcal{L}^*$ is the maximized value of the log-likelihood function provided by running the constrained optimization with log-likelihood function as the objective function. The value of ϵ is chosen by the econometrician.[19]

Note that I can take advantage of the linear form to greatly reduce the computation time and the number of constraints. Given any guess of $(\mu^s,\beta_0^s,\delta_0^s)$, we can construct $\{x_{js,t}\}_{j\in\mathcal{J}_J,t\in\mathcal{T}}$ to obtain the matrix of independent variable $\tilde{X}$. The solution $\hat{\alpha}$ that minimizes the GMM

[17] MPEC is not frequently used in trade. See Balistreri, Hillberry (2008); Lashkaripour, Lugovskyy (2018) for examples of applying MPEC method in trade.

[18] Other simple alternatives include using the simple or weighted sum of objective functions. I have tried both and they give similar results to the epsilon-method.

[19] The main challenge with this method is that the value of ϵ is chosen artfully by the econometrician. An ϵ too small will result in a problem with no feasible solution (as constraint not satisfied), and one too large renders the likelihood constraint useless. In my estimation, ϵ is 300 and $\mathcal{L}*$ ranges between 1400-1500 across industries.

objective function $g'Wg$ is the standard GMM estimator: $\hat{\alpha}_{gmm} = (\tilde{X}'ZWZ'\tilde{X})^{-1}\tilde{X}'ZWZ'y$ where W is the GMM weighting matrix.[20] Thus, the residual $\hat{\xi} = y - \tilde{X}\hat{\alpha}$ can be specified rather than solved for as in nonlinear demand system (e.g. in a random coefficient specification). This advantage reduces the number of constraints by almost half.

The optimization problem, written as a MPEC problem, is the following:

$$\min_{\beta_0,\delta_0,\mu,g} \quad g'Wg$$

subject to:

$$\begin{aligned}
c_1: &\quad x_{t+1} = \frac{\beta_t + \delta_t}{\beta_t + \delta_t + q_t} x_t + \frac{q_t\mu - r_t}{\mu(\beta_t + \delta_t + q_t)} \\
c_2: &\quad Z'\hat{\xi} = g \\
c_3: &\quad |\mathcal{L}(R_t|\hat{\theta}(\mu;\beta_0,\delta_0),Q_t) - \mathcal{L}^*| \leq \epsilon \\
c_4: &\quad \frac{\beta_0}{\beta_0 + \delta_0} \leq \mu
\end{aligned}$$

Constraint c_1 describes the motion of reputation; c_2 is the moment condition, c_3 specifies the likelihood function necessary to pin down μ, and c_4 guarantees that the initial reputation guess does not go beyond [0,1].

Constraints c_3 and c_4 restrict the values of learning parameters β_0, δ_0 and μ. In section 3.1, I mentioned that in the construction of c_3, exporters who have been in the U.S. market for fewer than 10 quarters are dropped. They are still included in the MPEC problem, entering in c_1, c_2 and the objective function. This means I still investigate how consumers respond to reputation of exporters who they don't learn much about. Countries that trade with the U.S. only temporarily are excluded from a constraint about learning parameters because they reveal little how consumers learn. If an exporter is not in the market ("no learning"), then the reputation stays unchanged.

3.4 Mapping from variables to data

Treating the United States as a representative consumer, we can map the variables on to data on an aggregate level. I_t maps on to the quarterly average household expenditure on the relevant consumption products. Within each HS6 category, price $p_{js,t}$ maps on to the unit value of the variety (a HS6-exporter pair) in that year; quantity $q_{js,t}$ maps on to the number of units, and $r_{js,t}$ maps on to the units of a HS6-exporter pair that is subject to at least one product recall. If no product s from country j is recalled within quarter t, then $r_{js,t} = 0$.

[20] I used the identity matrix as the weighting matrix in the estimation. I have also tried two-steps GMM,and $(Z'Z)^{-1}$, but both yield weighting matrices that are close to singular or badly scaled. Here I prioritize computation accuracy over asymptotic efficiency.

At the time a recall is issued, consumers receive information about certain product from an exporter. Assume that consumers consider the products imported from that country in a window around the recall to be problematic. In the baseline model, I assume that the window is three months after the recall occurs. For example, if a recall for Chinese toys happens in January 2008, all toys imported from China in January, February and March are considered affected by the event. Formally, $r_{js,t}$ can be calculated as:

$$r_{js,t} = \frac{\sum_{m \in t} Q_{js,t,m} \times \mathbb{1}(\mathrm{R}_{js,t,m} + \mathrm{R}_{js,t,m-1} + \mathrm{R}_{js,t,m-2} \neq 0)}{\sum_{m \in t} Q_{jk,t,m}}$$

where m is the subscript for months and t for quarter. If in a single month, multiple recalls for one variety is triggered, we still count the quantity only once in calculation of $r_{js,t}$. As in the previous example, if there is one recall in January and two in February, products imported from these two months are only counted once.

The market share I calculate in the data is the share of value:

$$S_{js,t} = \frac{p_{js,t} q_{js,t}}{\sum_{j'=1}^{J^s} p_{j's,t} q_{j's,t}} \tag{6}$$

For consistency of units, I calculate market share using value instead of units imported. The U.S. import data set reports two different units for some varieties. For example, in 1990, the port of Miami reported 1169 dozen, or 9096 kilograms (shipment weight), of men's suit jackets containing more than 36% wool imported from Colombia. Some exporters, however, only report one of the units. A common treatment in empirical analysis is to keep only the unit that exceeds the other in terms of numbers of units, but an inconvenience introduced by this treatment is that different exporters might use different units within one product market. This problem of "hidden varieties"–even finer differentiated varieties than the HS10 categories—–is a common problem in trade flow data. Computing market share in terms of the total value of imports—-a unique number for each entry reported in each year with unambiguous units-—-allows us to avoid the complication of units for reported quantity.

This problem is not a concern for the estimation of reputation. The fraction of product recalled is the key in computing reputation, so the unit of quantity is irrelevant. The units of recalled products $r_{js,t}$ and import $q_{js,t}$ for a variety are always the same.

By keeping parameters invariant across time and exporters, the framework assumes that consumers only "discriminate rationally". Namely, they differentiate exporters' products based only on the products' current reputations and the signals received in this period. The coefficient α_x^s governs the utility differentiation between a high quality and a low quality product s. The larger α_x^s is, the more consumers value a high quality product over a low one—in other words, consumers care about the quality of that product. As discussed in the introduction, in this empirical exercise, "quality" only concerns the safety of the product. For example, if α_x is higher in "toys and sports equipment" than "apparels," then we would conclude that consumers care more about safety of toys than clothes. Surely consumers want safe products in both categories, but the harm done to consumers by a toy with lead paint can be more severe than a battery that can overheat. μ is the probability of a recall if the product is of low quality. The arrival rate is determined by product characteristics and how consumers use them. When μ is high, we will consistently see frequent recalls for low reputation countries. When μ is low, fewer products are recalled per period and the variation relative to the mean of recall level is higher.

4 Data

4.1 Matching recall data to US import flows

To analyze the impact of informative signals on the market, I created a novel data set that links monthly U.S. import data from the Census to CPSC recall incidences from 1990 to 2009. I can observe the quantity and total value of import trade flow by trade partner, by month, and by HS10 product category. I then assigned a six-digit harmonized system code (HS6) to the products that are subject to recalls by reading through the descriptions of recall reports.

Although monthly trade data has HS10 level products, in the estimation I can only estimate reputation across exporter-HS6 varieties because recall events are only matched to HS6 level. The data appendix has a detailed discussion of the matching process and why it can only be reliably matched to HS6 level. The data is then aggregated to quarterly HS6 level, and a time period in the analysis will be a quarter henceforth. I need to aggregate monthly data to quarterly data because the computation of units affected by recalls requires one level of aggregation.[21].

The recall data set contains the date of the recall, the name and a brief description of the product, the types of hazards it brings, and its manufacturing countries.[22] In addition to the variables I scraped, the Consumer Product Safety Commission reports images of the products, remedies, the consumer contact, and manufacturers' or retailers' names. All incidences have a recall number, recall date, name, type, and description of the product and pictures. For more

[21] An alternative to aggregate over time is to aggregate over HS6 products. A major concern to that method is that by aggregating HS6 to, say, HS4, we are implicitly assuming that HS6 products within a HS4 category are perfectly substitutable. This is not true for some HS4 categories. For example, playing cards and game consoles are both HS6 products under category 9504, but they are not substitutable.

[22] In more recent recall events, the CPSC occasionally reported the price and units sold of the products recalled. The price and units sold are only available after October 1, 2010, so I did not use that information in this paper.

dated recall incidences, some information might be missing. A key piece of information from the CPSC is the manufacturing countries of the products and, as shown in table 1, from 1990 to 2009, only 74.3% of the reports recorded at least one manufacturing country. Each report contains a distinct recall ID. It is possible that in one report multiple products are recorded. That is less common in the entries from recent years, but is more likely for recall reports before 2000. In this case, if all the products recalled are from one HS6 category, I treat it as one incidence; otherwise, I record a separate incidence for each HS6 category included under a recall ID. A few reports record multiple exporters under one recall ID. In this case, I treat an incidence as a recall to each exporter.

The matching is done by reading the recall report title and description, so measurement error is possible. Most recall reports are matched to HS6 level, while some are matched to HS4 level. If a report cannot be matched even to HS4 level, it is categorized as "unmatched" and omitted from the data set. For consistency, I only used the incidences matched to HS6 level in this paper. The main difficulty in the matching process is caused by the difference in target audience of the Harmonized Tariff Schedule and the CPSC recall reports. The HTS schedule is designed for tariff purposes, so the users are customs officers and exporting firms. It specifies the types of goods and often the compositions of goods, which is a piece of information relevant for tariff purposes and known to the producers. The CPSC recall reports, however, provide a description of the end use and appearance of the product so consumers can immediately identify their purchase. For example, a harmonized tariff code will describe a product as "girl's cotton t-shirt, 90% cotton, 10% polyester" while the CPSC will describe it as "girl's red cotton t-shirt with Mickey Mouse". The data appendix provides a detailed example to illustrate why this issue limits the matching to HS6 level.

4.1.1 Macroeconomic Data

Besides the linked trade flow and recall data, I also need a measure for household budget constraint and the market share of the outside option. To measure a household's budget constraint on products in my data set, it is not desirable to examine U.S. household income or total expenditures since a large share of household expenditures will be on housing, food, transportation and utilities. Instead, I examine relevant categories of consumption goods expenditures by types of products table provided by the Bureau of Economic Analysis (BEA henceforth) using data from the Consumer Expenditure Surveys. The categories I examine are durable and non-durable goods expenditures, excluding food and beverage, motor vehicles, and gasoline.[23] I excluded those categories because the goods in them are not under the

[23]The categories I included are furnishings and durable household equipment, recreational goods and vehicles, other durable goods (like jewelry, books, luggage and phones), clothing and footwear, and other non-durable

administration of the CPSC, so they are irrelevant to this analysis. I used the BEA annual data, so the quarterly budget is a fourth of the yearly expenditure. All values are then discounted using Consumer Price Indexes from the Federal Reserve Bank of St. Louis, where 1982-1984 are the base years.

Discrete choice models allow consumers to have an outside option. Following Khandelwal's approach, the outside option here is to purchase from the United States. Using the annual production data reported in the NBER-CES Manufacturing Industry Database, the U.S. value of sales is calculated as the difference between the value of shipment and the U.S. export value in that year.

Table 2 summarizes the descriptive statistics of variables in the industries I estimated. [24]

4.2 Selecting products to estimate a learning model

The linked recall data set contains many products, but not all of them are suitable for estimating this learning model. There are two criteria that they need to satisfy: first, recalls are frequent enough that learning can plausibly happen, and second, the product is not durable.

The first criteria is straightforward: if a product only has a couple recalls over almost twenty years, then consumers do not have enough signals for learning to be meaningful. There will be almost no variation in reputation even if they are included in the estimation. Thus I keep only products that have at least 25 recall observations over the years, which is the 90th percentile of the 144 products that have at least one recall in the data set.[25] This cut leaves me 13 products.[26]

I limit the estimation to non-durable goods for both empirical and theoretical concerns. Among the 13 frequently recalled products, some varieties of have units values far exceed the average quarterly household expenditure, which is around \$1000 across the years. Ovens imported from United Kingdom, for example, have unit value exceeding \$1000 for 35 quarters. All of these products are expensive durable goods that consumers do not repeatedly purchase, at least not within a year or a quarter. Thus it is not appropriate to include them in the estimation of this particular learning model. I drop all the goods with a large fraction of high

goods (recreational items, household supplies, stationary). Some non-durable goods in "other non-durable goods" categories are also excluded. They are "pharmaceutical and other medical products" and "tobacco".

[24] I show the descriptive statistics of toys here, as the results in this industry will be presented in greater detail in Section 5 The rest of the industries will be discussed in Section 6

[25] Here, the recall observation is not an incidence, but a quarter-variety pair. If toys from Spain have recalls in January and March 2007, that will only count as one observation at 2007 Q1 in the product selection process. It will, however, count as two incidences, and it affects how we calculate the fraction of products recalled.

[26] They are toys, cotton sweaters, sweaters of man-made fabric, battery, lamps, hair dryers, ovens, cradles, stoves and ranges, snow mobile, baby trolley, and equipment for outdoor games.

unit value observations, and that are intuitively non-durable, which leaves me six products: toys,cotton sweaters, sweaters of man-made fabric, battery, lamps and hair dryers.

In the following section, I will present the parameter estimates and discussion of results for toys. Given that the model and data have variation across countries, products, and time, presenting results for one good helps us to focus on the cross-exporters and cross-time variation. The discussion illustrates mechanics and properties of the model. Once we have clarified the more subtle implications of the model, we will discuss the cross-product variation in next section. Toy is chosen as the example because it is the most frequently recalled product.[27] It also can cause serious health consequences in children, so consumers tend to value safety in this product.

5 Results in the Toys Industry

5.1 Reputation Formation

The update of reputation depends on learning parameters $[\mu, \beta_0, \delta_0]$ and the history of sales and recalls. Section 2 defines μ as the probability for a bad toy to be recalled. β_0 and δ_0 are initial values of distribution parameters that shape consumers' prior beliefs. Intuitively, β_0 is the units of toys ever recalled and δ_0 is the total units of un-recalled toys sold to the United States before April 1990.

I estimate the probability of recall μ using variation in units of recalls and quantity of imports. Intuitively, keeping the true fraction of unsafe products constant, if μ is close to 1, the model predicts more recalls with relatively small variance within each exporter, because exporters will see consistent recalls (or the absence of them). In the contrasting case when μ is close to 0, the model predicts few recalls with small variance because there is close to no recalls, and when μ is close to 0.5, some recalls but with larger variance. By fitting predicted recalls to actual recalls, μ can be identified as detailed in section 3.1.

The initial distribution parameters β_0 and δ_0 are selected using variation of recalls, quantity, and variation of constructed reputation. The ratio between β_0 and δ_0, can vertically shift the predicted reputation. The magnitudes of β_0 and δ_0 governs the impact of the first few periods of learning: intuitively, if β_0 and δ_0 are too small (relative to trade flows), the recalls in the first few periods will have a drastic impact on reputation, and if they are too large, reputation will not change much over 20 years.[28]

[27] Toys have 837 recall incidences over the years, followed by snowmobiles and golf carts, which have 136 recalls.

[28] Consider an example in which an exporter sells 1000 units to the U.S. every quarter, 10% of them are defects, and $mu = 1$. Supposing $\beta_0 = 3$ and $\delta_0 = 5$, after one period of update, we have $\beta_1 = 1003$ and $\delta_1 = 1005$. The

In addition to the variance of recalls, the model can distinguish between μ and initial distribution parameters β_0 or δ_0 by comparing the changes in reputations when a recall breaks out. Consider the cases of a low μ or a high β_0: both can lead to a lower initial value of reputation and shift the reputation downwards. Reputation is more responsive to recalls if μ is low because when bad products are unlikely to be recalled a recall become more alarming. To visualize how learning parameters change reputation, I plot estimated reputation with reputation constructed with manipulated learning parameters, keeping observed recalls and import quantity flows as given.[29] Figure 3 uses toys imported from Hong Kong to illustrate the vertical shift when β_0 or δ_0 are reduced by half, and figure 4 illustrates the change in reputation when μ decreases from 0.9115 to 0.6. We can see that reputation drops faster in periods with frequent recalls (say from the second quarter of 1990 to the first quarter of 1995) when μ is reduced than when β_0 is reduced, although the initial values are similar in two cases.

Panel 1 of table 3 presents the estimates for learning parameters. Panel 2 summarizes the reputation across exporters and time, constructed using the estimates in panel 1. Panel 3 lists periods of learning, the average number of quarters a country exports to the United States, initial reputation, reputation in the first period that is determined by estimates in panel 1, number of exporters ever selling to United States since 1990, and number of exporter-quarter pair in this industry. If all exporters stay through the 79 quarters in data set, we can hypothetically have $149 \times 79 = 11771$ observations. Instead, most exporters have only started exporting to the United States in recent years, so there are only 4344 exporter-quarter pairs in the data. After dropping some exporters who have only exported for a couple years to U.S., we have 3436 observations left to estimate $[\mu, \beta_0, \delta_0]$.

Panel 2 of table 3 shows the summary statistics of estimated reputation across time and exporters. In the last period, the exporters of toys with best reputations are Mexico and Canada, corresponding to the maximum 0.966 and 0.961; and the minimum corresponds to China. Canada has recalls in only two quarters of the 20 years in my observation, while China has at least one recall in 76 out of the 79 quarters. Both countries export to the United States in all periods, and they export in large quantity. Most exporters—127 out of 149—have never had a product recall by the CPSC. Canada's consistent presence in the U.S. market and large

change in reputation induced by recalls from the first period is $\frac{-100}{1000+5+3} \approx -0.099$, but in the second period will be $\frac{-100}{1000+1005+1003} \approx -0.033$. Changes in reputation vary widely if the initial guess β_0 and δ_0 are too small, and this variation does not reflect data patterns. In estimation I set the initial values for β_0 and δ_0 to have a comparable magnitude with trade flows, so the reputation variation in the first few periods are not too drastic.

[29] This simulation is distinct from the simulation I will discuss in section 5.4, although both exercises decrease μ. The purpose of this simulation is merely illustrating data pattern that I can use for identification, so recalls are taken as given. The simulation in section 5.4 also simulates recalls to understand the value of information for consumers.

exports make it stand out among the exporters who have always been safe.

In the preferred specification, I use the unit freight cost as the instrument for price. Table 4 shows that unit freight cost passes the "rule of thumb" test for instrument relevance (see Stock et al. 2002) for most non-durable goods, and it is strong for toys.[30] Exchange rate, though intuitively should be correlated with price, has a weak correlation. This is not surprising given how volatile exchange rate is over time and how big the variation is across currencies. Unit shipping cost and oil price times distance have the same channel: cost of transportation enters the CIF value in the import data set. When we include both, one of the two instruments will appear to be not-correlated, thus keeping only one is sufficient. Table 5 shows that including additional instruments does not change the results much. Note the different instrument specifications in table 5 are estimated from a two-step procedure instead of the one-step MPEC estimates reported in table 3. The two-step procedure takes the reputation constructed using learning parameters estimated in GMM, and runs regression 5. Although it is not the preferred specification since it cannot estimate all parameters simultaneously, it has significant speed advantage and I use it to illustrate alternative instrument specifications.[31] Comparing the results in column 5 in table 5 and table 3, we can see that the two-step procedure provides point estimates similar to the one-step MPEC estimates. Column 4 and 5 in table 5 show that adding additional instruments does not change the point estimates or F-statistics much.

5.2 Demand response to reputation

Panel 1 of table 3 displays α_x and α_0, the market share responses to reputation and log(budget-price). These two parameters reveal how sensitive consumers are to reputation and price. A positive coefficient for reputation implies that it is rewarding for exporters to maintain or aim for higher reputation, and it also means consumers are more concerned about reputation in this product.

The coefficient of reputation implies that the "reputation elasticity of market share" is 4.037 for toys.[32] If an average exporter of toys can increase reputation by 10%, it can expect to

[30] Stock et al. (2002) suggests that F-statistics<10 should raise concerns of weak instruments in the GMM estimation. Choosing an instrument that works for all industries is challenging, and unit freight cost is the best-performing instrument among those commonly used in the literature.

[31] In addition to time concerns, changing number of instruments is not trivial in the MatLab codes for the MPEC problem because I provide the Jacobian and Hessian matrices to speed up computation. Each additional instrument specification requires an different version of Jacobian and Hessian. The direction of change in demand coefficients should be the same between two-step procedures and the one-step MPEC, so to illustrate this point the two-step procedure suffice.

[32] Reputation elasticity of market share is the percentage change of market share induced by one percentage change of reputation. Use σ to denote the reputation elasticity, it is calculated as: $\sigma = \frac{d\ln s}{d\ln x} = \frac{d\ln s}{dx} \cdot \frac{dx}{d\ln x} =$

increase its market share by 40.37%. This is a somewhat big change, but given that reputation is history-dependent, it will take the average exporter many periods of safe presence in the U.S. market to achieve that.

To illustrate how long it will take an exporter to improve reputation, I take the reputation in the last period, and predict how long it will take for each exporter to increase reputation by 10% in two scenarios. The first scenario assumes that in each future quarter an exporter sells the same quantity into the United States, which is equal to the average quarterly quantity from the second quarter of 1990 to the last quarter of 2009. I run the reputation updating procedures with no recalls until the reputation reaches target level. An average large exporter of toys who is among the upper quartile in export quantities will need to have a safe presence for 35.9 years consecutively to improve its reputation by 10%. It will take even longer for small exporters because the information update is slow when consumers see few new units in the market. Even for the largest exporter of toys, China, catching up is difficult. It will take 754 quarters—that is 188.5 years—of flawless presence in the United States for its reputation to catch up with that of Mexico, the exporter currently enjoying the best reputation in the U.S. market.

In the second scenario, reputation growth rate polarizes when the simulation includes demand responses. I relax the assumption of sales volume in the previous simulation, allowing market share to change as reputation improves while fixing total units of sales. I simulate 10 million agents for 1000 quarters, each choosing an exporter in every period, and aggregate their choices to market shares. However, most exporting countries have market shares way smaller than 10^{-7}, so simulated market shares cannot match actual shares of these countries. To circumvent this problem, I focuses on 12 exporting countries who are the top 10% in export quantity.[33] The largest exporter, China, now takes only 273 quarters instead of 754 quarters to catch up with Mexico because its market share increases as reputation picks up. Canada and Mexico can improve their reputation to perfection in 757 and 94 quarters respectively, which is much longer compared to 34 and 29 in the simulation with the first scenario, as their market shares decline since their reputations cannot improve as fast as China's. All other countries in the simulation improve reputation by less than 10% within 1000 quarters, implying that it takes longer to improve reputation for most exporters when market share can change. Investing in quality inspection is more beneficial to large exporters, as they have initial and ongoing advantages in reputation growth.

$\alpha_x \cdot \dfrac{1}{1/x} = \alpha_x \cdot \bar{x}$, where $\bar{x}$ is the average reputation. The change in market share is relative to the U.S. market share since that is the outside option.

[33]Since the top 10% exporters can change from quarter to quarter, the final set is the union of all the top 10% countries in each period. These 12 countries together export 92.99% of foreign toys in the United States.

5.3 Discussion: Impact of a bad event

After establishing that recalls decreases an exporter's reputation, and that lower reputation leads to lower market share, we can quantify the impact of a recall event on market share. In this framework, the magnitude of impact for a recall event depends on the intensity of it, the size of exporter, history of the exporter's presence in U.S. market, and its current level of reputation. The marginal impact of recalling *one unit* of product at time t' on reputation in the periods following is:

$$\frac{\Delta x_{j,t+1}}{\Delta r_{j,t'}} = \begin{cases} \dfrac{1}{-\mu(\beta_0 + \delta_0 + \sum_{\tau=1}^{t} q_{j,\tau})} & \text{if} \quad t \geq t' \\ 0 & \text{otherwise} \end{cases}$$

Use $\Delta x_{j,t} \equiv x_{j,t} - x^0_{j,t}$ to denote the change of variable x in period t from $x^0_{j,t}$, the level it would be at had the recall not happen. Of course, when the CPSC issues a recall, not one toy train is recalled but an entire batch of it. Each recall event affects a number of products specific to the exporter, and its impact on reputation depends on the size of recall relative to the size of import from that exporter. Thus, to assess the impact of a recall, I fix the fraction of products that are recalled instead of units of products recalled. Here let us consider a recall event that will cause every unit of the product from country j be recalled. The difference in reputation induced by a recall that affects $r_{j,t'} = q_{j,t'}$ units of goods will change reputation by:

$$\Delta x_{j,t+1} = \frac{q_{j,t'}}{-\mu(\beta_0 + \delta_0 + \sum_{\tau=1}^{t} q_{j,\tau})} \quad \text{if} \quad t \geq t' \tag{7}$$

The impact of a recall event depends on when it happens and who it happens to. Another variation is the duration of impact: earlier recalls have a larger impact as it will influence—though with diminishing effect—all the periods following. We also expect that a recall happening in a year of large export volume will have a strong impact, as more units are affected.

Taking quantities and parameter estimates of μ, β_0, δ_0, and α_x as given, I use the change in reputation from equation 7 to calculate the marginal impact on relative market share $\tilde{s}$ as displayed in equation 8.[34]

$$\frac{\tilde{s}_{j,t+1} - \tilde{s}^0_{j,t+1}}{\tilde{s}^0_{j,t+1}} = \exp\left(\frac{\alpha_x q_{j,t'}}{-\mu(\beta_0 + \delta_0 + \sum_{\tau=1}^{t} q_{j,\tau})}\right) - 1 \tag{8}$$

[34]Substituting equation 7 into equation 5, we can get equation 8: $\Delta \ln \tilde{s}_{j,t+1} = \alpha_x \Delta x_{j,t+1}$ where $\tilde{s}_{j,t+1} \equiv \frac{s_{j,t+1}}{s_{USA,t+1}}$. Note that $\Delta \ln \tilde{s}_{j,t+1} = \ln\left(\frac{\tilde{s}_{j,t+1}}{\tilde{s}^0_{j,t+1}}\right)$. We can then write the percentage deviation of relative market share from $\tilde{s}^0_{j,t+1}$ caused by a recall in period t' ($t \geq t'$)as the following:

$$\frac{\tilde{s}_{j,t+1} - \tilde{s}^0_{j,t+1}}{\tilde{s}^0_{j,t+1}} = \exp(\alpha_x \Delta x_{j,t+1}) - 1 = \exp\left(\frac{\alpha_x q_{j,t'}}{-\mu(\beta_0 + \delta_0 + \sum_{\tau=1}^{t} q_{j,\tau})}\right) - 1$$

The impact of recall events on market shares varies across exporters and across quarters. Each possible recall event dampens that exporter's reputation in all quarters following, and the impact of a recall is calculated as the discounted sum of impacts in all future quarters.[35] I calculated the impact by quarters of recall occurrence, but for visual clarity I sum quarterly impacts into annual impacts and plot the spectrum of impacts across exporters for each year.

Panel 1 in figure 5 sums up the variation of recall impact by year of occurrence for all exporters. Each box-plot is a distribution of percentage change of own-country market share for recalls that happened in the corresponding year in the x-axis. The strong negative impact in year 1998-2000 is driven by a large quantity of import in those years. Across all years, an average exporter will lose 2.15% of its market share for a recall event that is severe enough to affect every unit of import during the year, and for most exporters, their loss does not exceed 7% of their market share. This means that the magnitude of impact from a single recall event is not detrimental, even though the impact persists for all following periods. Consumers seem more lenient compared to what Freedman et al. (2012) finds because the agent who holds reputation in my context is a country instead of a firm or an individual player.[36] Prices vary widely across exporters. Consumers are price-sensitive, so they are willing to accept some risk of getting a bad product if it is cheap enough.

Panel 2 in figure 5 illustrates the loss in trade values for an average exporter of toys in a recall event. The average loss in a year is 2.437 million dollars. The pattern over time is similar to the pattern in market share changes: large import quantity drives large changes. Note that in early 1990s the total import quantity and value are low, so early recalls do not have as big an impact in trade value as in market share.

5.4 Discussion: Quantifying the Value of Information

Every year, the Consumer Product Safety Commission submits a budget request to the Congress. For example, the budget request for fiscal year 2019 is 123.5 million dollars. Thus from a policy maker's perspective, it is meaningful to ask how important a quality inspection institution like the CPSC is to domestic consumers. The model answers this question from an information perspective.

Consider two scenarios, one in which the inspection institution can catch and recall unsafe products more effectively than the other. Under the more effective scenario ("high inspection accuracy"), if a product is unsafe, it will be caught with 90% chance while in the other scenario ("low inspection accuracy"), that probability is 50%. Note that a low μ does not mean noisier

[35] The quarterly discount factor is 0.995, so the annual discount rate is 0.98.

[36] Freedman et al. (2012) finds that unit sales of a category of toys from a manufacturer (firm) decreases by 38.9% on average if it is recalled in 2007.

signals: recalls still only signal unsafe products, but the signals are rarer. I measure welfare changes using compensating variation, that is, the changes in income to make consumers indifferent between having high and low inspection accuracy. I assume that in both scenarios, the size of market and underlying fraction of unsafe products are the same. Let x^L denote the reputation in the low accuracy scenario and x^H in high accuracy scenario. The compensating variation $\text{cv}_{s,t}$ satisfies:

$$\alpha_0 \log(I_t - p_{j^*s,t}) + \alpha_x^s x_{j^*s,t}^H + \eta_{j^*} + \psi_t = \alpha_0 \log(I_t - p_{j's,t} + \text{cv}_{s,t}) + \alpha_x^s x_{j's,t}^L + \eta_{j'} + \psi_t$$

Note that, here, j^* is the exporter that consumer chooses in high inspection accuracy scenario, and j' is the exporter chosen in other scenario. j^* and j' need not be the same. I assume that when the quality of signal is low, consumers are aware of it and incorporate that knowledge in learning.

I take the underlying fraction of unsafe products as given, and simulate the recall events and consumer learning under high inspection accuracy ($\mu = 0.9$) and low accuracy ($\mu = 0.5$), and I compute the compensating variation for consumers of toys.[37] To simulate recall events, I assume the last period reputation is the best proxy for the true unobserved fraction of bad products. Taking quantity imported in the United States as given, the number of bad products $L_{js,t}$ is the product of reputation estimates in the last period and the quantity of imports. Each unit of bad product has probability μ of being recalled, so the total number of products recalled roughly follows a normal distribution with mean $\mu L_{js,t}$ and variance $\mu(1-\mu)L_{js,t}$. After generating number of products recalled for each exporter in each quarter, I can run the reputation updating following equation 2 to estimate $x_{j's,t}^L$ or $x_{j's,t}^H$ under each scenario.

The simulation generates 12,000 agents with individual preferences drawn from an Extreme Value distribution, and it provides two measures of welfare: the total compensating variation for the U.S. market and the average compensating variation for each purchase.[38] For each exporters j in each quarter t, a set of simulated agents choose their products (that set can be empty). The average compensating variation within the consumers for an exporter, multiplying the total units of products from the corresponding exporter in period t, gives us the simulated compensating variation for consumers buying from country j in time t. The sum of each exporter yields the total compensating variation. Average compensating variation is calculated as the total compensating variation divided by the units of product s imported into the U.S. market in period t, which is equivalent to the average of simulated agents' compensating

[37] I am not aware of any empirical work that specifies the effectiveness of CPSC recalls, so there is no obvious benchmark for this exercise. I pick the high μ as it is close to the estimated value of μ in toys, and low μ to have an equal chance between recall and no recall.

[38] Extreme value distribution takes location parameter $\mu = 0$ and scale parameter $\sigma = 1$.

variation weighted using the import share of the countries agents choose to buy from.[39] Total compensating variation is driven by both the change in average compensating variation—a channel that reveals the impact of information—and the changes in demand.[40] While total compensating variation highlights the magnitude of impact, the average compensating variation excludes the impact of import quantity, so it can better reveal the model mechanisms.

The welfare loss per purchase averages around \$0.87 over time when inspection accuracy is low.[41] If we consider the volume of purchase in toys, however, the total welfare loss can average 695 million dollars per quarter. Panel 1 of figure 6 shows the total compensating variation for toys. The welfare loss from lower inspection accuracy comes from lower mean utility when μ is low and also higher chance of landing a unsafe product.

Two Mechanisms of Utility Changes

To fully understand the sources of welfare differences under two scenarios, figure 6 decomposes the two channels through which welfare changes, and I call them the "mean value difference" and "defect surprise." Using Utility(H)' to denote the maximized consumer utility under high μ scenario, after realization of the product quality, and Utility(L) to denote the mean utility under low μ (weak inspection) scenario, the following equation describes the decomposition.

$$\overbrace{\text{Utility(H)'} - \text{Utility(L)'}}^{\text{Utility gain from having higher }\mu} = \underbrace{\overbrace{[\text{Utility(H)'} - \text{Utility(H)}]}^{\text{Utility loss from recall (high }\mu\text{ scenario)}} - \overbrace{[\text{Utility(L)'} - \text{Utility(L)}]}^{\text{Utility loss from recall (low }\mu\text{ scenario)}}}_{\text{"defect surprise"}}$$
$$- \underbrace{[\text{Utility(H)} - \text{Utility(L)}]}_{\text{Mean utility differences}}$$

When the probability of bad products getting recalls is low, consumers evaluate exporters differently and have a more pessimistic reputation assessment. The expected utility differs because prices will be different for the consumers who choose another exporter in the alternative scenario, and reputation changes for all consumers. In addition, consumers who get a unsafe product will take a utility reduction after revelation, and I call this damage "defect surprise." A

[39]The equations to calculate total and average compensating variation from the simulation are the following. Let $\text{cv}(j^i)_{i,t}$ denote the compensating variation for individual i who chose exporter j^i to purchase from in period t, and q_{j^i} denote the quantity imported from exporter j^i. Total $\text{CV}_t = \dfrac{1}{12000}\displaystyle\sum_{i=1}^{12000} \text{cv}(j^i)_{i,t} \times q_{j^i}$ and Average $\text{CV}_t = \dfrac{\text{Total CV}_t}{\sum_{j\in J} q_{j,t}}$

[40]The total quantity demanded is not explicitly modeled in this framework as I focus on changes in market share, the demand *relative to* your competitors given the number of consumers.

[41]All dollars are converted into 1982-1984 dollars using CPI, and later quarters are discounted using discount factor 0.995.

positive "defect surprise" suggests that utility loss from recall is less damaging when inspection is more effective. Under the weak inspection scenario, consumers will observe fewer recalls but treat each one with greater caution because they know the probability of recall is lower. It will take them longer to approach a more accurate estimate of true fraction of defect products. Thus under weak inspection consumers will be surprised with a defect product more, which incurs a cost illustrated as a curve *above* horizontal axis in figure 6.

Figure 7 illustrates that when μ is low, the reputation estimates are lower because consumers have a more pessimistic prior, but eventually the reputation will catch up and approach the "true fraction of good products" specified in the simulation. Exactly how long it will take to converge back to the true fraction, however, depends on the quantity of trade flows. Figure 8 shows that the reputations of China are similar under both scenarios because China is a large exporter throughout the years, but for Mexico the discrepancy remains large till the late 1990s when the quantity of toys sold to the U.S. increases. Thus lower inspection accuracy decreases reputations for all exporters, but the damages are more severe and long-lasting for small exporters.

As a result, the mean value difference can have ambiguous impact. A positive mean value difference, illustrated as when the curve is above horizontal axis, means that the expected utility is higher when inspection is strong. However, weaker inspection can sometimes increase consumers' expected utility, because the marginal consumers may now buy from a large exporter. If large exporters happen to sell cheaper products and the reduction in consumer expenditures compensates the reduction in reputation, then the saving can lead to higher mean utility. Comparing two scenarios, the marginal consumer "switch to" larger exporters because their reputations reduce less compared to smaller exporters.[42] As we can see from figure 6 though, usually the reduction of reputation creates a loss in utility that far exceeds the price differences.

The difference between "defect surprise" and "mean value difference", which is the area between two curves, is the costs to having less effective inspection calculated in compensating variation. In the rare case when the benefit of cost-saving outweighs the higher risk of getting a defect, it is possible that better inspection is not welfare-improving. However, the simulation suggests that it is unlikely, and under most scenarios better inspection improves consumer welfare.

[42]Since these two hypothetical scenarios cannot co-exist, there are no actual switchers. The marginal consumers "switch" in the sense that they will choose differently under the alternative scenario.

Market share changes after a decrease in μ

Simulation also reveals that smaller exporters benefit more from a highly effective inspection institution. Figure 9 compares market shares when inspection accuracy is high and low. All exporters lose market shares when μ is low because now purchasing from any exporter is perceived to be riskier and consumers prefer the outside option. After several periods however, the market share recovers and the lowest reputation exporter—China—even have a small gain in market share towards the second half of the observed periods. This is seemingly surprising, until we realize that the low reputation exporter (China) also happens to be the largest exporter. A downward shock in inspection effectiveness hurt reputation of all exporters, but larger exporters recover faster. Given the large import volume from China, the gap between and after the shock closes much earlier for China than other exporters, so it actually gains a temporary "advantage": not from "better" reputation, but from resilience to information quality shock.

6 Results across Industries

In the previous section, I use toys as an example to illustrate model mechanisms and what they can do in terms of welfare and counterfactual analysis. This section introduces estimation results for other products, revealing heterogeneity in consumers' concern of safety across products. The results carry interesting policy implications for any exporter improving quality but have limited resources and for domestic institutes like the CPSC who may need to budget quality inspection expenditures across types of products.

Table 6 shows the difference across products in term of consumers' preferences for reputation. Column 1 and 2 show the coefficients from the MPEC estimation, and column 3 and 4 show the corresponding market share elasticities. Sweaters of man-made fabric is the product that consumers have the strongest preference for safety, with a market share elasticity of reputation of 5.07, followed by cotton sweaters and toys. Unsurprisingly, sweaters of different materials have similar demand elasticities. Similar to toys, improving reputation by 10% can lead to a big increase in market share for exporters of sweaters, by 42.47% and 50.68% respectively. Compared to toys and clothes, consumers only have a weak preference for a safe battery, and do not seem to care whether lamps and hair dryers post a safety hazard.

The differences in types of hazards post by these products can explain some of the differences between consumer preferences. Table 7 lists the most frequent hazards for each type of products, and we can see that the most frequent hazards for toys and sweaters either can be fatal (choking and strangulation) or can cause long-term distress for users (lead paint). For example, a pullover sweater presents a "choking hazard" when a stitched-on flower can fall off, and children may accidentally swallow it. It is worth noting that the majority of the apparels recalled by the CPSC are clothes for children, although the harmonized system code category can only describe the product as “Shirts; men's or boys' ”. Consumers' preference may not only reflect the types of hazards, but also to whom hazards may occur: the same hazard can be far more damaging when it happens to a vulnerable child, which can explain the larger coefficient estimates for toys and clothes.

6.1 Discussion: cross-industry differences in recall impact and welfare implication

Recalls to which product are the most harmful to exporters? Which group of consumers need accurate quality inspection more than others? To answer these questions, I perform the same exercises described in section 5.3 and 5.4.

Table 8 shows that products that consumers care more about—toys and children's clothes—have bigger per recall event impact on average. A recall event is again defined as an event that can affect all units of products that exporter sells to the United State that quarter. The costs depends on the size of the exporter at that time, so it varies across time and exporters. An event as described will cost an average exporter of cotton sweaters 3.34 million dollars in value of exports, almost twenty times as much as the costs for an average exporter of battery. Note that an actual recall event will rarely last for a year, or cover 100% of products imported from an exporter, with but a few exceptions. The lead paint scandal in 2007, for example, causes ongoing recalls for Chinese toys for almost two years.

Although the demand elasticities of reputation are similar for toys and sweaters, the market share impact of a recall event is larger for sweaters than for toys. This is driven by both the differences in μ and different market structures across products. The estimate for μ is smaller for sweaters than for toys. Reputations are more responsive to recalls when μ is small, so are market shares. When I replace the estimated μ for sweaters with estimated μ for toys in a simulation, the market share response decreases to 7% for cotton sweaters and 6.48% for sweaters of man-made fabric. In addition, the market for toys is more concentrated than the market for sweaters. There are more exporters of sweaters than toys consistently, and summary statistics presented in table 2 show that the leading exporters of sweaters do not have as dominating a market share (44.2% at highest) as the leader of toy exporters (89.6% at highest).[43] Market concentration affects market share responsiveness mostly through the parameters β_0 and δ_0 because they capture the average past history of recalls and sales. Intuitively, in a highly concentrated market, some exporters may be way smaller than the

[43] Figure A.3 in Appendix A illustrates the number of exporters of toys and sweaters.

average, and when new recalls occur, the additional information is diluted by the relatively big denominator that contains β_0 and δ_0. As a result, their reputation will not be as responsive to recalls.

Figure 6 illustrates the large differences among the welfare changes for consumers of toys, sweaters and battery when the probability of a bad product being recalled decreases from 90% to 50%. Average total compensation in a quarter is 695 million dollars for toys, but it is only 0.088 million, 0.013 million and 0.16 million dollars for cotton sweater, sweaters of man-made fabric, and battery respectively. The difference is driven by both the difference in per unit purchase welfare change, and in the market size. Average welfare loss per unit of purchase is 87 cents for toys, and 0.078, 0.028, and 0.71 cents for cotton sweaters, sweaters of man-made fabric and battery respectively. The magnitude of change in toys market is over 1000 times bigger than that in sweaters, and it is around 100 times the size of the change in battery. It is expected that the welfare impact for battery consumers is small, since the demand elasticity for reputation is only about 10% of that of toys and sweaters. The welfare impact for sweater consumers is low most because consumers appear to be price-sensitive in sweaters market. That means a relatively small increase in total income can compensate for the loss of utility from receiving a defect product. Another interesting pattern we can see in sweaters of man-made fabric panel is that sometimes having weaker inspection does not harm consumer welfare. This is an example of our discussion in section 5.4, in which consumers are price-sensitive enough to value price reduction over reputation reduction.

These results suggest that quality inspection institutions like the CPSC do benefit consumers, but to a different extent depending on the types of products. If importers or exporters decide to invest in quality inspection, they should prioritize products primarily used by children, since consumers seem to have strong preferences for safe products in these categories. However, reputation improvement can take decades even for large exporters.[44] For most exporters of products used by children, improving reputation can increase their market share. That may not be the case for exporters of other consumption goods, so exporters may have weaker incentives to invest in quality control, and choose to compete through lower prices.

7 Conclusion

This paper analyzes the effect of an exporter's reputation on import trade flows. It defines an exporter's reputation as the expected probability of drawing a high quality product in

[44] It is generally hard for small exporters to improve reputation, but it is especially hard for small exporters who used to be large. More developed Asian exporters (like Hong Kong and South Korea) have displayed this pattern for products like toys.

a market; and it adopts a framework in which consumers Bayesian update their belief of exporters in a product market. This paper tackles the challenge of identifying intangible and unobserved reputation in two ways: constructing a data set in which I can see shocks that affect reputation, and modeling channels in which reputation affects consumers' decisions. Compared to other empirical papers studying the reputation of sellers, this analysis reveals a variation of impacts across a broad set of products. The model in this paper can be generalized to estimate consumers learning of any signals in trade, for example, how the market reacts to a scandal that is widely cover in traditional and social media, like the Vokswagan diesel emission scandal.

This paper is a step towards understanding the role of consumers' learning in international trade. There are at least three directions of future research. First,this model uses Bayesian learning–a type of perfect learning–with perfect memory, and this is an idealistic assumption of the market. I can generalize this model to incorporate imperfect memory models, and explore how reputation dynamic changes. Second, this paper focuses on estimating the learning dynamic for goods that are purchased frequently. Durable goods likely have a different information acquisition dynamic that we can explore. Third, this model abstracts away from firms' decision on investing in quality improvement. Given that reputation matters for some products, incorporating the producer's decision is a natural next step.

Tables and Figures

Table 1: Recalls reported: manufacturing countries and number of matches

	Number of Reports	**Fraction of total**
Matched to HS6 Code	3217	0.617
Matched to HS4 Code	619	0.119
Does not report manufacturing countries	1342	0.257
Cannot match, other	36	0.007
Total	**5214**	**1**

Note: This table reports the match quality of recall incidences to trade flows from 1990-2009. Source of recall incidences is the CPSC recall database.

Table 2: Summary Statistics

Variables	Statistics	Toys	Sweaters[a]	Sweaters[b]	Battery	Lamps	Hair Dryers
Market share	Mean	0.00154	0.00979	0.00952	0.0153	0.00764	0.0536
	Median	0.00011	0.0011	0.0004	0.00073	0.0002	0.00189
	Max	0.896	0.403	0.442	0.467	0.406	0.623
	Min	8.12×10^{-8}	8.89×10^{-8}	1.82×10^{-7}	4.28×10^{-6}	3.09×10^{-6}	1.67×10^{-5}
Price	Mean	28.18	27.45	32.91	40	68.65	24.36
	Median	6.35	18.06	19.79	14.07	29.26	16.32
	Max	467.65	122.35	149.33	419.17	481.02	145.51
	Min	0.02	1.2	1.08	0.07	0.49	1.51
Quantity (in millions)	Mean	14.5	0.896	0.497	1.15	0.372	0.518
	Median	0.0299	0.7168	0.0156	0.0138	0.00337	0.00786
	Max	1700	59.1	31.8	47.1	15.5	6.807
	Min	2×10^{-6}	3×10^{-6}	2×10^{-6}	5×10^{-6}	3×10^{-6}	1.8×10^{-5}
Value of Trade (in millions)	Mean	32.7	13.5	7.69	5.29	3.25	3.797
	Median	0.218	1.16	0.32	0.181	0.0785	0.133
	Max	3060	1040	107	195	200	53.9
	Min	2.57×10^{-4}	2.51×10^{-4}	2.52×10^{-4}	1.256×10^{-3}	1.26×10^{-3}	1.294×10^{-3}
Units of Recall (in millions)	Mean	0.833	0.06	0.0166	0.239	0.0921	0.061
	Median	0	0	0	0	0	0
	Max	1570	59.1	29.5	47.1	21.7	5.87
	Min	0	0	0	0	0	0
Ratio of Recall	Mean	0.0316	0.00522	0.00368	0.00982	0.00584	0.0189
	Median	0	0	0	0	0	0
	Max	1	1	1	1	1	1
	Min	0	0	0	0	0	0
US market share	Mean	0.128	0.104	0.149	0.503	0.638	0.364
	Median	0.118	0.0491	0.126	0.416	0.629	0.357
	Max	0.312	0.346	0.458	0.963	0.784	0.61
	Min	0	0.00342	0.0147	0.279	0.526	0.272

Note: a: Sweaters made of cotton, HS6=611020. b: Sweaters made of man-made fabric, HS6=611030.
Source of trade data is the monthly U.S. Census import data. Recalls come from the CPSC recall database. U.S. manufacturing data comes from NBER-CES data set. All summary statistics are reported from the quarterly data set aggregated from monthly data. Each variable means: 1) market share calculated from import values. 2) row reports unit value of import. 3) quantity imported to the U.S. in the unit that reports a larger number of quantity. 4) value of trade in current USD. 5) quantity of recalled products in the same unit as import quantity in 2). 6) ratio of recall to import quantity. 7) U.S. market share.

Table 3: Parameter Estimates for toys

Parameter Estimates			
Description	*Parameter*	*Estimate*	*S.E.*
Recall probability given product is low quality	μ	0.9115	(0.0242)
Sum of recalled units before 1990 (millions)	β_0	82.75	(0.0689)
Sum of units of sale before 1990 (millions)	δ_0	145.9	(6.0766)
Preference for Reputation	α_x	6.433	(1.17)
Coefficient of log(budget-price)	α_0	16.16	(0.168)

Descriptive Statistics of Reputation in the Last Period				
	Mean	*Std. Dev.*	*Min*	*Max*
All Countries	0.6235	0.0783	0.0751	0.9656
Highest reputation quartile	0.6987	0.0989	0.6065	0.9656
Lowest reputation quartile	0.5888	0.0868	0.0751	0.6030
Conditions of Learning				
Periods of Learning	29.154	28.02	1	79
Initial Reputation	0.6030	-	-	-
Number of Countries	149	-	-	-
Number of Observations	3436	-	-	-

Note: μ is robust to different initial guesses. I chose 15 guesses spacing equally between 0.1 and 1: all return the same estimate. Initial guess for β_0 is 10 times the average units of recalled products; and for δ_0 10 times the average units of goods sold.

Table 4: First Stage OLS Regression: $\ln(I_{i,t} - p_{js,t})$ on unit freight cost

HS6	**Products**	**Coeff.**	**S.E.**	**F-stat**
950300	Toys	-0.004149	0.000162	655.36
611020	Sweater, cotton	-0.000937	0.000290	10.43
611030	Sweater, man-made fabri	-0.000942	0.000369	6.5
850780	Battery	-0.004215	0.000333	160.02
940520	Lamps	-0.004514	0.000220	422.3
851631	Hair dryers	-0.000931	0.002348	0.16

Note: Regressing log(expenditure-price) on unit costs.

Table 5: Logit Estimates of Demand, Toys only, All Exporters

	$\ln(s) - \ln(s_0)$				
Reputation	4.594	5.159	2.148	5.081	5.076
	(0.840)	(0.528)	(0.883)	(0.530)	(0.530)
log(expenditure-price)	27.39	56.66	39.26	64.41	64.94
	(0.905)	(1.565)	(1.392)	(2.409)	(2.405)
Two way FE	No	Yes	No	Yes	Yes
IV: Unit Transportation Cost	No	No	Yes	Yes	Yes
IV: Exchange Rate	No	No	Yes	No	Yes
IV: Oil Price×Distance	No	No	Yes	No	Yes
Observations	4344	4344	4344	4320	4320
F	536.7	1251.5	472.5	1236.7	1235.8

Note: Standard errors in parentheses. Coefficients are estimated using the two-steps procedure in which reputation is constructed using learning parameters estimated from one-step MPEC procedure, and then I run a logit demand regression taking constructed reputation as given. Observations different in the last two columns as 24 singleton groups are dropped.

Table 6: Preference estimates across industries, non-durable goods

Products	**Coefficient**		**Elasticity**		**Obs**
	Reputation	**log(expenditure-price)**	**Reputation**	**Price**	
Toys	6.433	16.155	4.037	-0.422	4344
	(1.174)	(0.168)			
Sweater, cotton	4.936	111.255	4.247	-2.763	6983
	(0.266)	(0.032)			
Sweater,	5.960	94.211	5.068	-2.789	6064
man-made fabric	(0.236)	(0.079)			
Battery	0.805	21.838	0.543	-0.787	2216
	(1.269)	(0.310)			
Lamps	-0.421	15.738	-0.158	-1.0103	3097
	(1.112)	(0.035)			
Hair dryers	-0.109	61.747	-0.0481	-1.3606	934
	(0.818)	(0.022)			

Note: Standard error in parentheses. Standard errors are GMM standard errors calculated with identity weighting matrix.

Table 7: Top five most frequent hazards for different products

Toys hazard	**Percentage**	**Sweaters hazard**	**Percentage**
Choking	52.57	Strangulation	53.98
Lead	19.32	Fire; fire-related burn	23.01
Electrocution/Electric Shock	6.72	Choking	20.35
Laceration	4.65	Entanglement	1.77
Fire; fire-related burn	3.06	Entrapment	0.88
Hair dryer hazard	**Percentage**	**Lamps hazard**	**Percentage**
Fire; fire-related burn	43.69	Fire; fire-related burn	40.61
Electrocution/Electric Shock	35.44	Electrocution/Electric Shock	34.55
Burn - Not Fire-Related	16.99	Collapse	9.09
Choking	1.46	Laceration	6.67
Fall	1.46	Burn - Not Fire-Related	3.03

Source: the CPSC recall database.

Table 8: Average impact of a recall event, per quarter

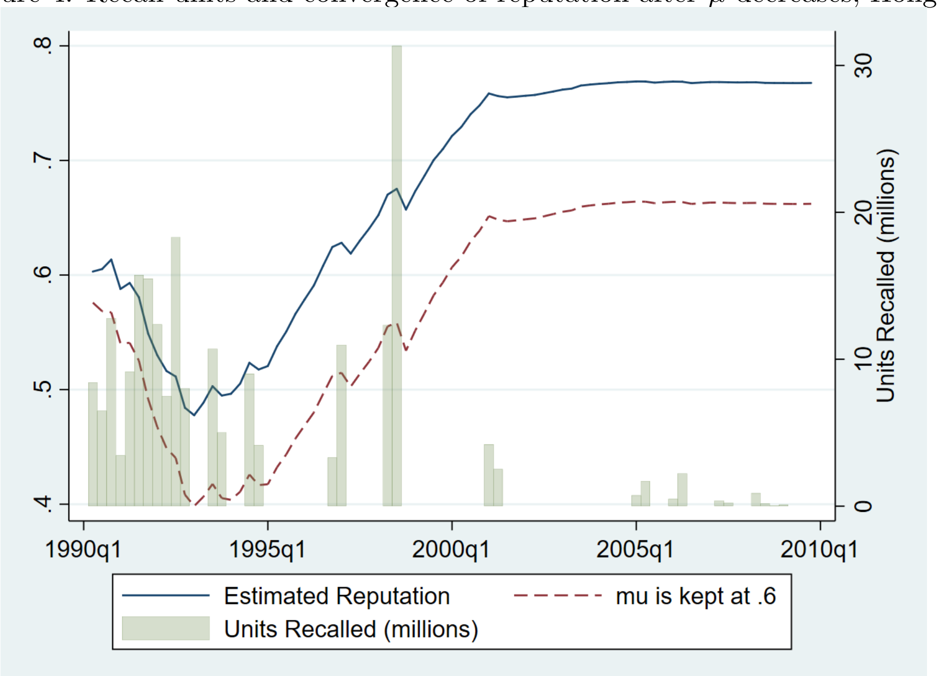
Products	Value (millions)	Market Share (%)
Toys	-2.437	-2.15
Sweater, cotton	-3.34	-16.47
Sweater, man-made fabric	-2.943	-26.29
Battery	-0.177	-1.43

Note: recall event is define as an incidence that affects 100% of the goods imported from that exporter in the period. Average across exporters and across time. Quarterly discount factor is 0.995. All values are normalized to 1982-1984 US dollars using CPI.

Figure 3: Reputation changes under different β_0 and δ_0, Hong Kong toys

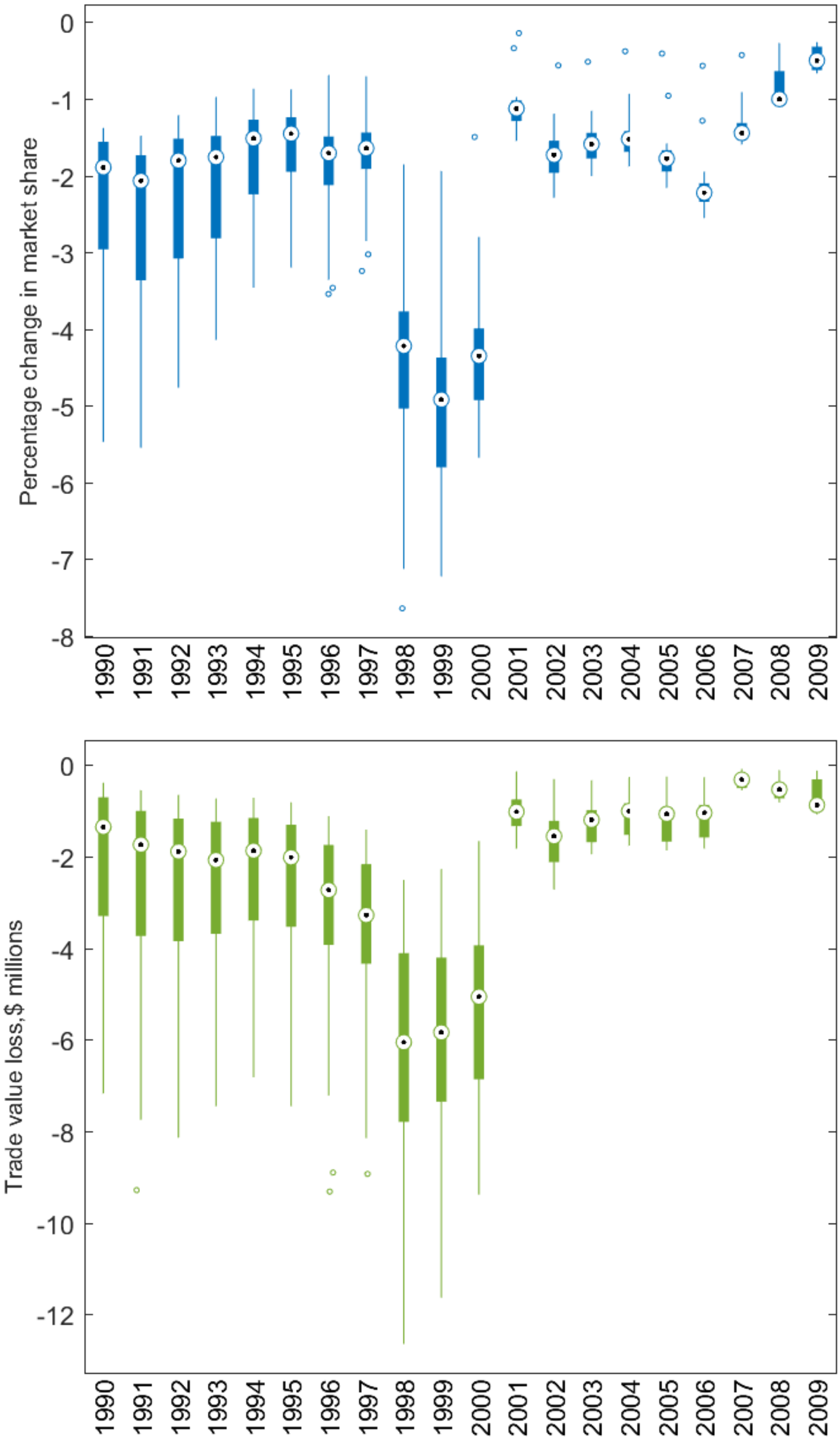
Note: this graph illustrates how reputation changes with recalls, and how estimated reputation changes when β_0 and δ_0 changes. Recall data from the CPSC recall data set. Learning parameters for construction of recall data set are reported in table 3.

Figure 4: Recall units and convergence of reputation after μ decreases, Hong Kong



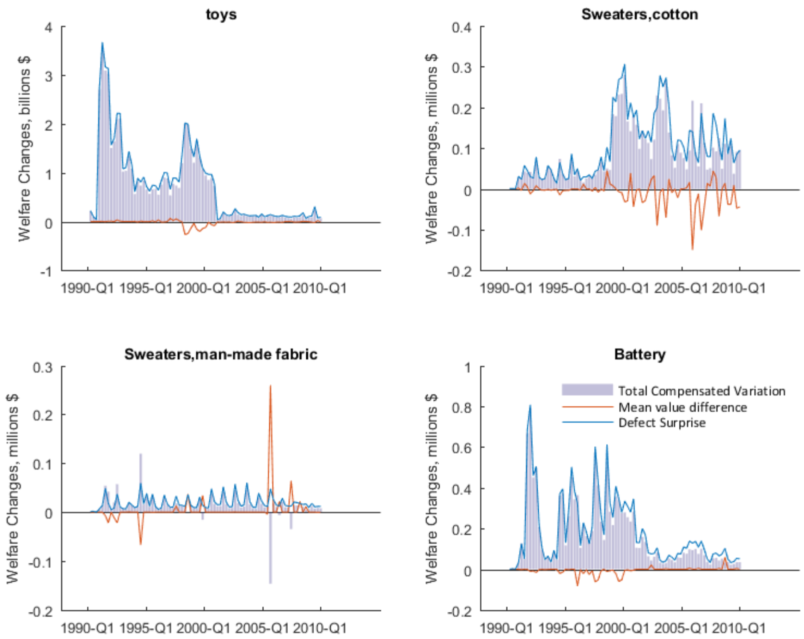
Note: this graph illustrates how estimated reputation changes when μ changes.Recall data from the CPSC recall data set. Learning parameters for construction of recall data set are reported in table 3.

Figure 5: Impact from a recall event for toys



Note: This figure plots the impact of a recall event (defined as an event that affect 100% of products imported in the year). The x-axis marks the time of occurrence for the event, and the recall will affect all periods following. For each recall, its impact varies across countries and across future periods. The tops and bottoms of each "box" are the 25th and 75th percentiles of the recall impact, respectively. The dot in box marks median, and the line (whisker) marks full range of observations. Hollow dot marks outliers. All values in 1982-84 dollars and discounted using quarterly discount rate 0.998.

Figure 6: Total compensating variation in 1982-84 dollars



Note: this figure plots the welfare loss when $\mu = 0.5$ instead of $\mu = 0.9$. All in 1982-84 dollars, discounted using quarterly discount rate 0.998.

Figure 7: Simulated reputation changes, Hong Kong toys

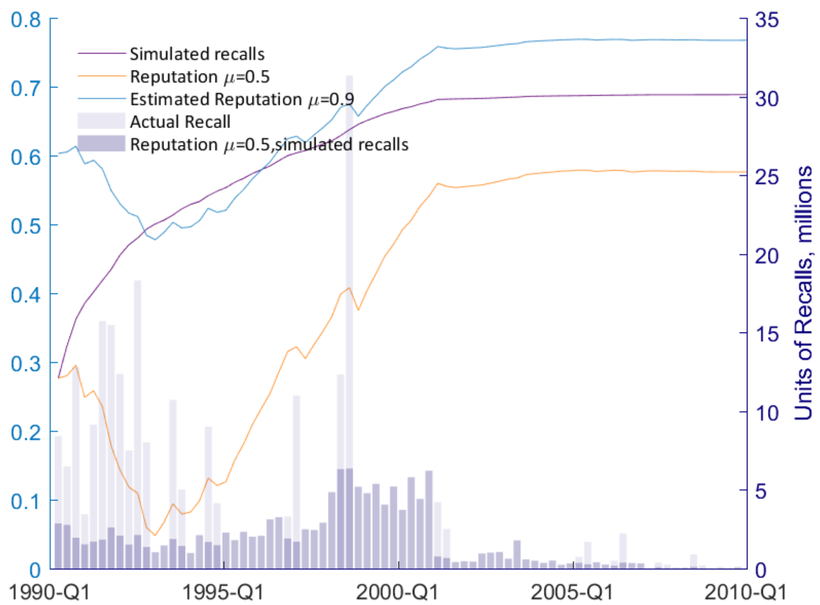


Figure 8: Import units of toys and convergence of reputation after μ decreases

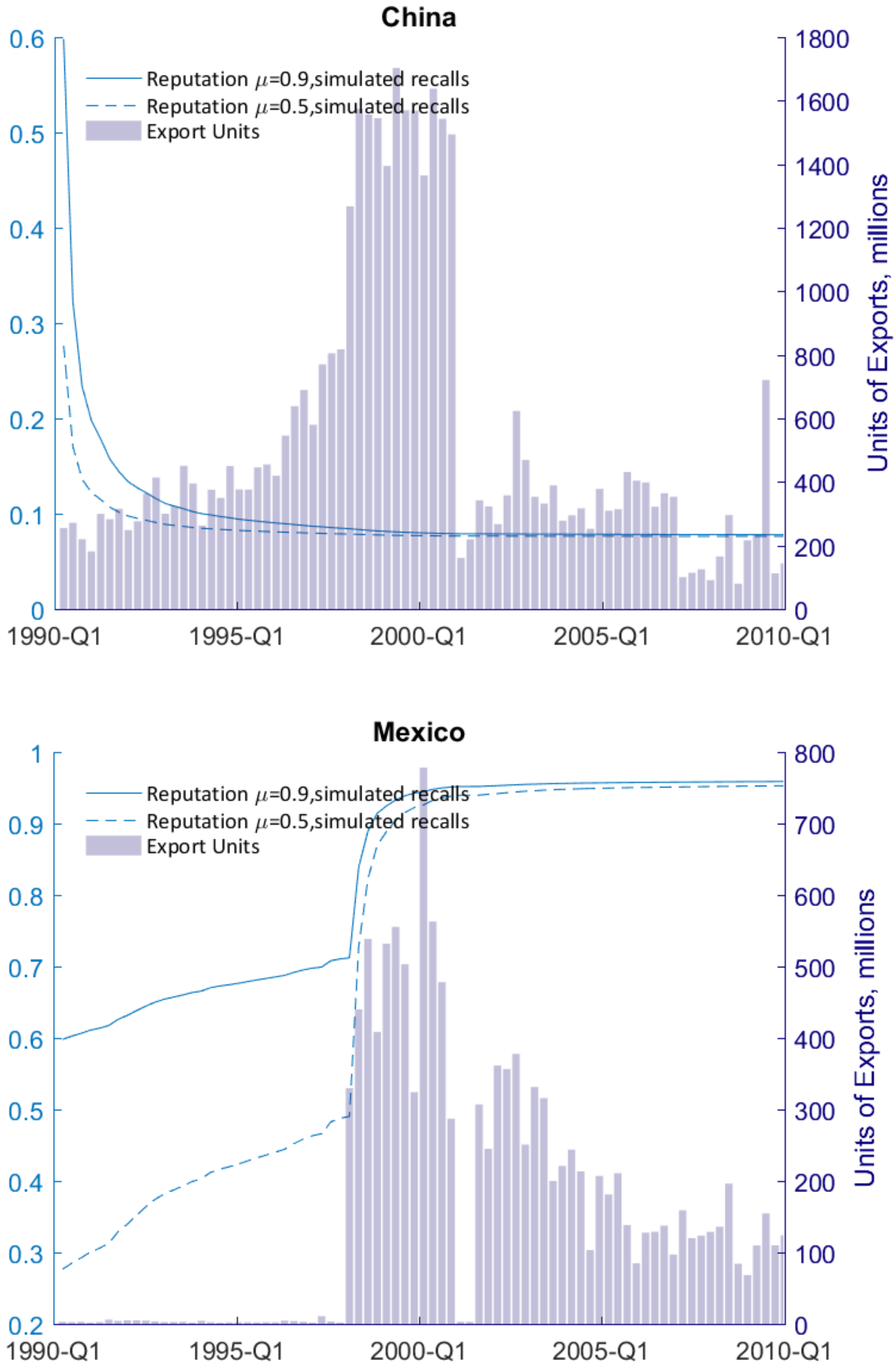
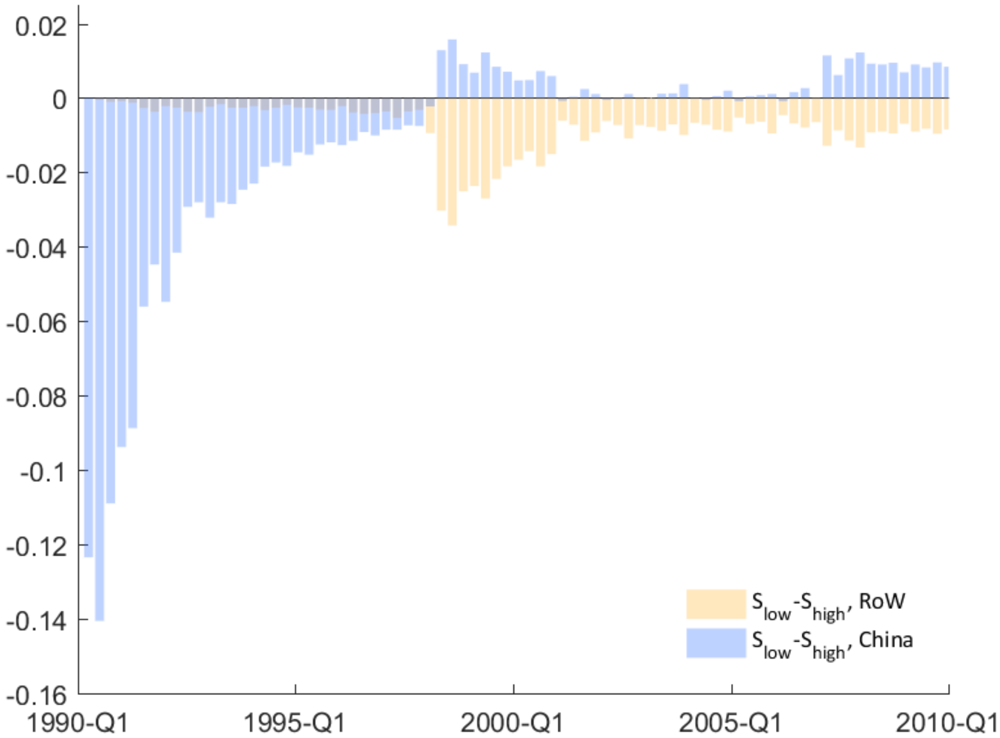


Figure 9: Simulated market share differences between $\mu = 0.9$ and $\mu = 0.5$, toys



Note: This figure plots the difference between market share if μ decreases from 0.9 to 0.5 in the first period, comparing the case of China and rest of the exporters (RoW). The United States is not included in this plot.

References

Albornoz Facundo, Pardo Héctor F Calvo, Corcos Gregory, Ornelas Emanuel. Sequential exporting // Journal of International Economics. 2012. 88, 1. 17–31.

Balistreri Edward J, Hillberry Russell H. The gravity model: An illustration of structural estimation as calibration // Economic Inquiry. 2008. 46, 4. 511–527.

Bar-Isaac Heski. Reputation and Survival: learning in a dynamic signalling model // The Review of Economic Studies. 2003. 70, 2. 231–251.

Bar-Isaac Heski, Tadelis Steven, others . Seller reputation // Foundations and Trends® in Microeconomics. 2008. 4, 4. 273–351.

Berry Steven T. Estimating Discrete-Choice Models of Product Differentiation // Source: The RAND Journal of Economics Journal of Economics. 1994. 25, 2. 242–262.

Board Simon, Vehn Moritz Meyer-ter. Reputation for quality // Econometrica. 2013. 81, 6. 2381–2462.

Bolton Patrick, Harris Christopher. Strategic experimentation // Econometrica. 1999. 67, 2. 349–374.

Bond Eric W. International trade with uncertain product quality // Southern Economic Journal. 1984. 196–207.

Chen Maggie X, Wu Min, others . The Value of Reputation in Trade: Evidence from Alibaba. 2016.

Chisik Richard. Reputational comparative advantage and multinational enterprise // Economic Inquiry. 2002. 40, 4. 582–596.

Commission The Consumer Product Safety. Products Under the Jurisdiction of Other Federal Agencies and Federal Links. 2016. Accessed: 2016-12-1.

Commission United States International Trade. HTS Archive. 1987-2012. Accessed: 2016-11-26.

Commission United States International Trade. By Chapter of HTS :2011-01-01 - Basic, Official Harmonized Tariff Schedule of the United States Annotated. 2011. Accessed: 2016-11-26.

Creane Anthony. Ignorance is bliss as trade policy // Review of International Economics. 1998. 6, 4. 616–624.

Creane Anthony, Miyagiwa Kaz. Information and disclosure in strategic trade policy // Journal of International Economics. 2008. 75, 1. 229–244.

Daalhuis Adri B. Olde. NIST Digital Library of Mathematical Functions. 2018.

Donnenfeld Shabtai. Intra-industry Trade and Imperfect Information about Product Quality // European Economic Review. 1986a. 30, 401417.

Donnenfeld Shabtai. Intra-industry trade and imperfect information about product quality // European Economic Review. 1986b. 30, 2. 401–417.

Donnenfeld Shabtai, Mayer Wolfgang. The quality of export products and optimal trade policy // International Economic Review. 1987. 159–174.

Donnenfeld Shabtai, Weber Shlomo, Ben-Zion Uri. Import controls under imperfect information // Journal of International Economics. 1985.

Dubé Jean-Pierre, Fox Jeremy T, Su Che-Lin. Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation // Econometrica. 2012. 80, 5. 2231–2267.

Eaton David H. Valuing Information: Evidence from Guitar Auctions on eBay // Journal of Applied Economics and Policy. 2005. 24, 1. 1–19. Name - eBay Inc; Copyright - Copyright Western Kentucky University 2005; Document feature - Equations; ; Last updated - 2016-06-11; SubjectsTermNotLitGenreText - United States–US.

Eaton Jonathan, Eslava Marcela, Krizan Cornell J, Kugler Maurice, Tybout James. A search and learning model of export dynamics // Penn State University mimeo. 2009.

Falvey Rodney E. Trade, quality reputations and commercial policy // International Economic Review. 1989. 607–622.

Falvey Rodney E, Kierzkowski Henryk. Product quality, intra-industry trade and (im) perfect competition. 1984.

Freedman Seth, Kearney Melissa, Lederman Mara. Product recalls, imperfect information, and spillover effects: Lessons from the consumer response to the 2007 toy recalls // Review of Economics and Statistics. 2012. 94, 2. 499–516.

Grundke Robert, Moser Christoph. Hidden protectionism? Evidence from non-tariff barriers to trade in the United States // Working Paper. 2014.

Haimes YV. On a bicriterion formulation of the problems of integrated system identification and system optimization // IEEE transactions on systems, man, and cybernetics. 1971. 1, 3. 296–297.

Holloway Isaac R. Learning via sequential market entry: Evidence from international releases of US movies // Journal of International Economics. 2017. 104. 104–121.

Houser Daniel, Wooders John. Reputation in auctions: Theory, and evidence from eBay // Journal of Economics & Management Strategy. 2006. 15, 2. 353–369.

Jacod Jean, Protter Philip. Probability essentials. 2004.

Khandelwal Amit. The long and short (of) quality ladders // The Review of Economic Studies. 2010. 77, 4. 1450–1476.

Scale Economies and the Structure of Trade and Industrial Policy. // . 2018.

Livingston Jeffrey A. How valuable is a good reputation? A sample selection model of internet auctions // Review of Economics and Statistics. 2005. 87, 3. 453–465.

Mayzlin Dina, Dover Yaniv, Chevalier Judith. Promotional reviews: An empirical investigation of online review manipulation // The American Economic Review. 2014. 104, 8. 2421–2455.

Newey Whitney K. Efficient Instrumental Variables Estimation of Nonlinear Models // Econometrica. 1990. 58, 4. 809–837.

Pierce JR, Schott PK. VConcording US Harmonized System Codes Over Time // Journal of Official Statistics. 2009.

Potoski Matthew, Prakash Aseem. Information asymmetries as trade barriers: ISO 9000 increases international commerce // Journal of Policy Analysis and Management. 2009. 28, 2. 221–238.

Rauch James E, Watson Joel. Starting small in an unfamiliar environment // International Journal of industrial organization. 2003. 21, 7. 1021–1042.

Reynaert Mathias, Verboven Frank. Improving the performance of random coefficients demand models: the role of optimal instruments // Journal of Econometrics. 2014. 179, 1. 83–98.

Rothschild Michael. A two-armed bandit theory of market pricing // Journal of Economic Theory. 1974. 9, 2. 185–202.

Saeedi Maryam. Reputation and adverse selection, theory and evidence from eBay // Working Paper. 2014.

Stock James H, Wright Jonathan H, Yogo Motohiro. A survey of weak instruments and weak identification in generalized method of moments // Journal of Business & Economic Statistics. 2002. 20, 4. 518–529.

Swann Peter, Temple Paul, Shurmer Mark. Standards and trade performance: the UK experience // The Economic Journal. 1996. 1297–1313.

Wu Min. The Value of Reputation in Trade: Evidence from Alibaba // Work. 2016.

Appendix A Additional Figures

Figure A.1: Correlation between quantity and recall incidences

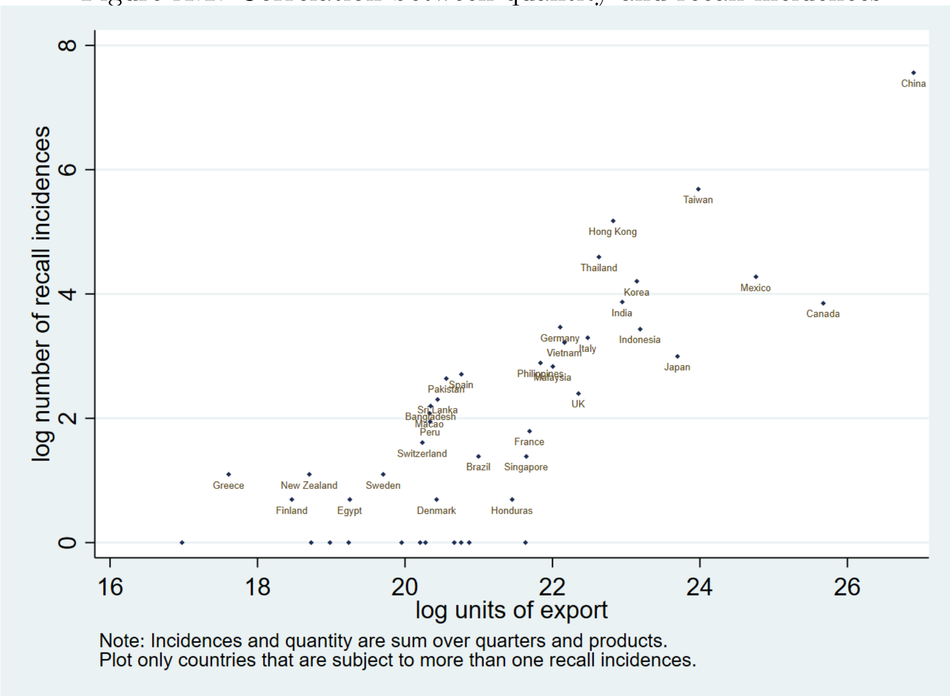


Figure A.2: Market share and recalls of toys from Hong Kong

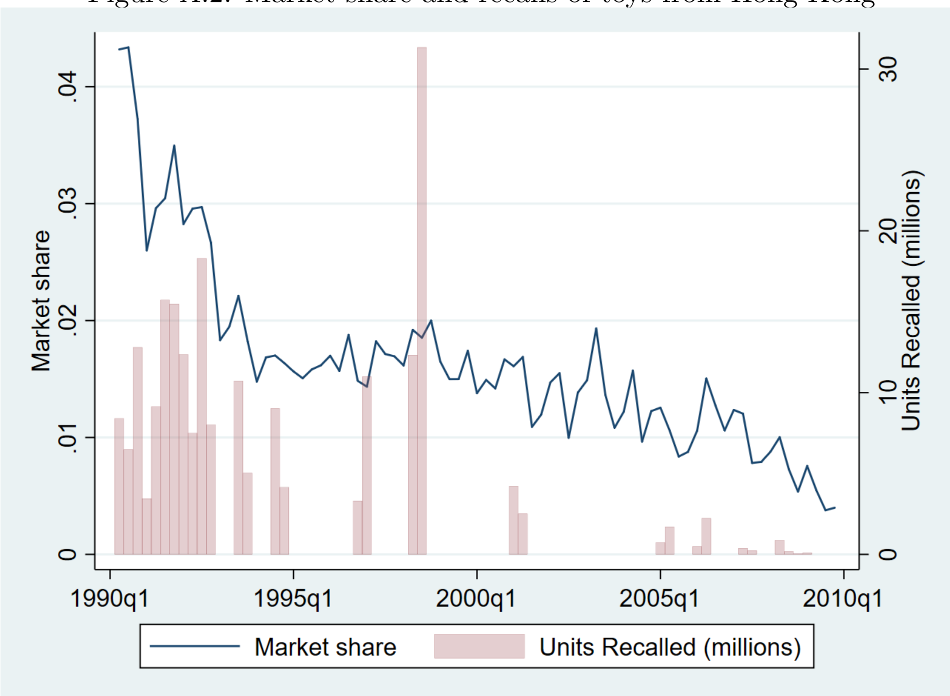
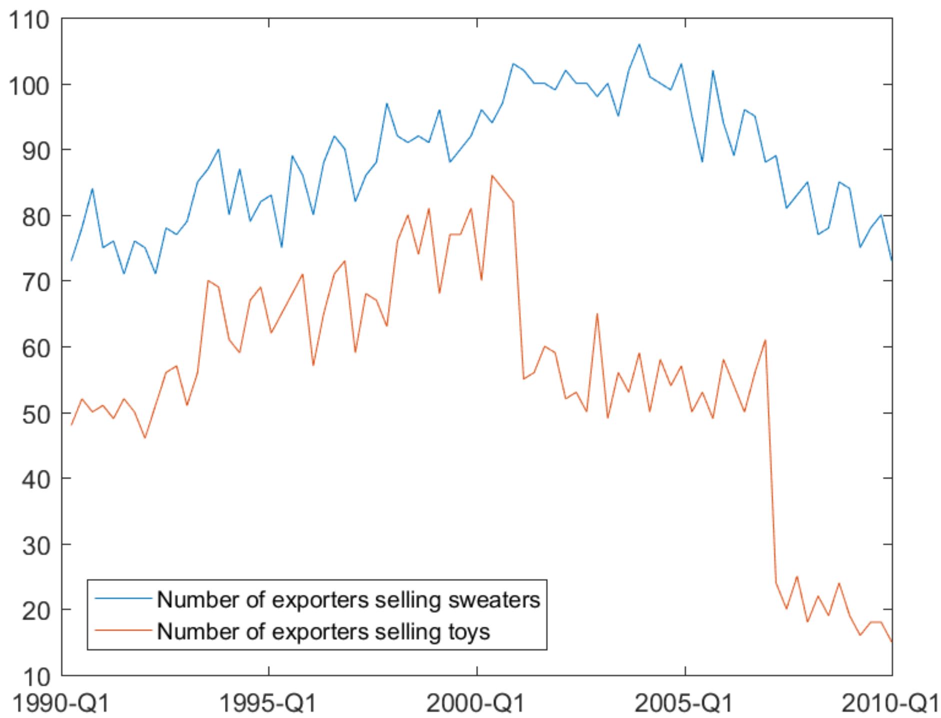


Figure A.3: Correlation between quantity and recall incidences



Appendix B Data Appendix

From Peter Schott's data set, I can also know additional information about the way of transportation (air, vessel or containerized vessel) and the cost of transportation, but I am not using these information in the analysis.

The set of information provided by Consumer Product Safety Commission (CPSC) does not perfectly overlap with the set of information used to describe a HTS category. Take an example of a recall occurred on November 3, 2011:

> ***Boy Scouts of America Recalls Cub Scout Wind Tech Jackets Due to Strangulation Hazard***
>
> *Description:*
>
> *This recall includes the blue Cub Scout Wind Tech jacket sold in youth sizes. The jackets are nylon with a polyester lining, long-sleeve, with a full zipper front and a Cub Scout wolf head emblem embroidered on the upper left chest. SKU numbers 73291, 73292, and 73293 are printed on the hang-tag that is attached to the jacket at retail.*

This report is categorized, according to HTS schedule 2011 (Commission 2011), under HS code 620193–"*Anoraks (including ski-jackets), wind-cheaters, wind-jackets and similar articles; men's or boys', of man-made fibres*". The title specifies that it is a boy's jacket, which pins it down to the category of men and boys' outwear (6201); and the information "nylon with polyester lining" in the description allows me to further narrow it to the category "boy's jacket with man-made fibres" (620193). To further refine this particular category however, I will need information on the composition which is not available in the recall data scraped. For example, the eight digits HS code 62019325 is described as "*......Containing 36 percent or more by weight of wool or fine animal hair*".

In the previous example, from the description I can still gather enough information to assign a six-digit HS code to the report. In some cases, the match is impossible without further research on the products. Here's another example from a recall report filed in 2005:

> *The candle holder is a Christmas decoration designed to hold a tealight candle. The candle holder includes three figures (penguin, moose, snowman) dressed in red and green sweaters, scarves and hats, roasting marshmallows on a stick over a small fire. Model numbers 4-01-427, 231279-4 and UPC code 90000 08741 are printed on the bottom of the candle holder.*

In this case, as shown in table B, candle holders of different materials belong to distinct HS2 industries. Thus it is impossible to assign the report into any category when the material of the candle holder is not specified. Among the 5214 reports from year 1989-2012, [blank] are not categorized for the lack of relevant information and I drop them out of the sample.

Table 9: Candleholder Materials and Corresponding HS8 codes, 2016 HTS

Material	Corresponding HS8 code
Glass	70139900
Wood	44209090
Metal	83062990
Ceramic	69120090

Another challenge in data mapping is the change of Harmonized Tariff Schedule over time. This problem cannot be ignored because HTS changed multiple times over the twenty-three years the data covers and some categories that went through major changes–toys, for example–made up a big proportion of the recalls occurred. I used 2002 HTS schedule as the main reference to construct a preliminary matching, then I used the harmonized system codes concordance over time provided by Pierce and Schott (Pierce, Schott 2009) to identify categories and spots that have undergone changes. Pierce and Schott provided concordance from 1989-2004, and adjust the matching manually by checking HTS schedule year by year (Commission 1987-2012). From 2004 onward, I adjust the matching by checking HTS schedule Archive. This process is finished within a reasonable time because my data set contains only 35 HS2 industries. Further more, I double checked matches in the top 5 industries in 1989-2004 using the HTS schedule on USITC website. Recall intensity is quite top concentrated: among those industries, 12 out of 35 industries have less than ten reports, 14 have 10-100 reports and only 9 have over 100 reports. The top five industries with the most recall reports consist of 74.8% of the recall reports, thus by performing the double check on the top industries, I made sure that a majority of the reports are matched to a correct HS6 code.

It is not surprising that the sample contains only 35 industries. Consumer product safety commission recalls a wide range of consumer products, but compared to the range of intermediate and final goods United States imports, it is a much smaller set. Also, some recalls are not issued by CPSC and will not show up in my data set. For example, food, cosmetics and drugs recalls will be under the administration of Food and Drug Administration. Automobiles, trucks, motorcycles and parts of them will be recalled by National Highway Traffic Safety Administration(Commission 2016). I kept only industries that have at least one recall from 1989-2012 in the merged data: all other industries are excluded because they are out of the administrative responsibility of CPSC or a recall is so rare it did not happen in the twenty three years. The latter case is quite unlikely; and although by reading a description of CPSC on range of products under their jurisdiction I can infer a set of industries that might be relevant, this process may introduce unnecessary measurement error.

Appendix C Mathematical Appendix

C.1 Derive the Dynamic Reputation Update Equation

The updating of reputation follows the Bayes rule. When choose a Beta distribution $\mathcal{B}(\beta_0, \delta_0)$ as the initial prior for $\mu(1-\theta)$, reputation updating follows:

$$\rho(r, \theta) = \rho(r|\theta) \times \rho(\theta)$$

The posterior density is:

$$\rho(\theta|r) = \frac{\rho(r|\theta) \times \rho(\theta)}{\rho(r)}$$

In the baseline model, when we choose a Beta distribution $\mathcal{B}(\underline{\beta}, \underline{\delta})$ as the prior distribution, after one period of learning, the updated joint density $\rho(r, \theta)$ still follows a Beta distribution:

$$\rho(r, \theta) \propto \gamma^{\overline{\beta}-1}(1-\gamma)^{\overline{\delta}-1}$$

and the distribution parameters update through: $\overline{\beta} = \underline{\beta} + r$ and $\overline{\delta} = \underline{\delta} + q - r$.

The mean of a Beta distribution $\mathcal{B}(\beta, \delta)$ is $\dfrac{\beta}{\beta + \delta}$. Thus the expectation of γ after one period of observation is updated as the following:

$$\begin{aligned}
\mathbb{E}[\gamma|r] &= \frac{\overline{\beta}}{\overline{\beta} + \overline{\delta}} \\
&= \frac{\underline{\beta} + r}{\underline{\beta} + \underline{\delta} + q} \\
&= \frac{r}{\underline{\beta} + \underline{\delta} + q} + \frac{\underline{\beta} + \underline{\delta}}{\underline{\beta} + \underline{\delta} + q}\, \mathbb{E}[\gamma] \qquad (9)
\end{aligned}$$

And from the definition of γ, substitute in $\theta = 1 - \dfrac{\gamma}{\mu}$, we can rewrite equation 9 as an equation of $\mathbb{E}[\theta|r]$ and $\mathbb{E}[\theta]$:

$$\begin{aligned}
\mathbb{E}[\theta|r] &= 1 - \frac{\mathbb{E}[\gamma|r]}{\mu} \\
&= 1 - \frac{1}{\mu}\left[\frac{r}{\underline{\beta} + \underline{\delta} + q} + \frac{\underline{\beta} + \underline{\delta}}{\underline{\beta} + \underline{\delta} + q}[\mu(1 - \mathbb{E}[\theta])]\right]
\end{aligned}$$

Given β_0 and δ_0 as the initial parameter values for the Beta distribution, in period t, the updated distribution parameters are $\beta_t = \beta_0 + \sum_{\tau=1}^{t-1} r_\tau$ and $\delta_t = \delta_0 + \sum_{\tau=1}^{t-1} q_\tau - \sum_{\tau=1}^{t-1} r_\tau$. Thus the reputation evolves from period t to period $t+1$ following:

$$\begin{aligned}
x_{t+1} &= E(\theta|r_t) \\
&= \frac{\beta_t + \delta_t}{\beta_t + \delta_t + q_t} x_t + \frac{q_t \mu - r_t}{\mu(\beta_t + \delta_t + q_t)}
\end{aligned}$$

C.2 Discussion: using truncated Beta as the prior distribution

In last section, I shown the reputation updating process derived when the prior distribution is a standard Beta distribution. An alternative assumption that fits the model intuition better is a truncated Beta distribution that limits the support for γ to be $[0,\ \mu]$. Here, I discuss a truncated Beta instead of a μ-scaled Beta or generalized Beta because the latter two are not conjugate priors of the Bernoulli trials, although they have cleaner functional form for the first moment. I will show how a truncated Beta is also a conjugate prior and how its mean is close to the mean of standard Beta when β_t and δ_t are large.

Conjugacy

Suppose we choose a truncated prior

$$p_B(\gamma|\underline{\beta},\underline{\delta}) \propto \gamma^{\underline{\beta}-1}(1-\gamma)^{\underline{\delta}-1}\mathbb{1}(0 \leq \gamma < \mu)$$

The likelihood function is

$$\mathcal{L}(\gamma) \propto \gamma^{r}(1-\gamma)^{q-r}$$

Thus the posterior distribution is:

$$\begin{aligned} p(\gamma|y) &\propto \gamma^{\underline{\beta}-1}(1-\gamma)^{\underline{\delta}-1}\mathbb{1}(0 \leq \gamma < \mu)\gamma^{r}(1-\gamma)^{q-r} \\ &\propto \gamma^{\underline{\beta}-1+r}(1-\gamma)^{\underline{\delta}-1+q-r}\mathbb{1}(0 \leq \gamma < \mu) \\ &\propto \gamma^{\overline{\beta}-1}(1-\gamma)^{\overline{\delta}-1}\mathbb{1}(0 \leq \gamma < \mu) \end{aligned}$$

where $\overline{\beta} = \underline{\beta} + r$ and $\overline{\delta} = \underline{\delta} + q - r$.

First moment

The p.d.f. corresponding to the truncated prior is:

$$f(\beta,\delta) = \frac{\gamma^{\beta-1}(1-\gamma)^{\delta-1}\mathbb{1}(0 \leq \gamma < \mu)}{B(\beta,\delta)F(\mu)}$$

where F is the c.d.f. of the Beta distribution. $F(\mu) = \displaystyle\int_0^\mu \frac{x^{\beta-1}(1-x)^{\delta-1}}{B(\beta,\delta)}dx$. For notational simplicity, write the numerator in the form of a incomplete Beta function

$$B(\mu,\beta,\delta) = \int_0^\mu x^{\beta-1}(1-x)^{\delta-1}dx$$

. Thus $F(\mu) = \dfrac{B(\mu,\beta,\delta)}{B(\beta,\delta)}$ The p.d.f. of the truncated Beta can be written as:

$$f(\beta,\delta) = \frac{\gamma^{\beta-1}(1-\gamma)^{\delta-1}\mathbb{1}(0 \leq \gamma < \mu)}{B(\mu,\beta,\delta)}$$

The expectation of γ in each time period is thus:

$$\begin{aligned} \mathbb{E}[\gamma_t] &= \int_0^\mu \frac{\gamma \cdot \gamma^{\beta_t-1}(1-\gamma)^{\delta-1}}{B(\mu,\beta,\delta)}d\gamma \\ &= \frac{\int_0^\mu \gamma^{\beta_t}(1-\gamma)^{\delta-1}}{B(\mu,\beta,\delta)}d\gamma \\ &= \frac{B(\mu,\beta_t+1,\delta_t)}{B(\mu,\beta_t,\delta_t)} \end{aligned}$$

Thus the mean of γ_t is a ratio of two incomplete Beta functions. For notational simplicity, let us drop the time subscript in the following proofs. Variables and data are all product-country-time specific.

$$
\begin{aligned}
&\frac{\beta}{\beta+\delta}B(\mu,\beta,\delta) - B(\mu,\beta+1,\delta) \\
&= \frac{1}{\beta+\delta}\left[\int_0^{\mu} \beta x^{\beta-1}(1-x)^{\delta-1}dx - (\beta+\delta)\int_0^{\mu} x^{\beta}(1-x)^{\delta-1}dx\right] \\
&= \frac{1}{\beta+\delta}\int_0^{\mu} x^{\beta-1}(1-x)^{\delta-1}\left[\beta-(\beta+\delta)x\right]dx \\
&= \frac{1}{\beta+\delta}\int_0^{\mu} x^{\beta-1}(1-x)^{\delta-1}\left[\beta(1-x)-\delta x\right]dx \\
&= \frac{1}{\beta+\delta}\int_0^{\mu} \left[\beta x^{\beta-1}(1-x)^{\delta} - \delta x^{\beta}(1-x)^{\delta-1}\right]dx \\
&= \frac{1}{\beta+\delta}\left(x^{\beta}(1-x)^{\delta}\Big|_0^{\mu}\right) \\
&= \frac{1}{\beta+\delta}\left[\mu^{\beta}(1-\mu)^{\delta}\right] \qquad (10)
\end{aligned}
$$

Rearrange the results from equation 10 into ratio form:

$$
\begin{aligned}
&\frac{\beta}{\beta+\delta}B(\mu,\beta,\delta) - B(\mu,\beta+1,\delta) = \frac{1}{\beta+\delta}\left[\mu^{\beta}(1-\mu)^{\delta}\right] \\
\Longrightarrow\ & B(\mu,\beta+1,\delta) = \frac{1}{\beta+\delta}\left[\beta B(\mu,\beta,\delta) - \mu^{\beta}(1-\mu)^{\delta}\right] \\
\Longrightarrow\ & \frac{B(\mu,\beta+1,\delta)}{B(\mu,\beta,\delta)} = \frac{\beta}{\beta+\delta} - \frac{\mu^{\beta}(1-\mu)^{\delta}}{(\beta+\delta)B(\mu,\beta,\delta)} \qquad (11)
\end{aligned}
$$

The incomplete Beta function has the following property according to (Daalhuis 2018):

$$
\begin{aligned}
B(\mu,\beta,\delta) &= \frac{\mu^{\beta}(1-\mu)^{\delta}}{\beta}\tilde{F}(\beta+\delta,1;\beta+1;\mu) \\
&= \frac{\mu^{\beta}(1-\mu)^{\delta}}{\beta}\left(\sum_{s=0}^{\infty}\frac{(\beta+\delta)_s\cdot\mu^s}{(\beta+1)_s\cdot s!}\right) \\
&= \frac{\mu^{\beta}(1-\mu)^{\delta}}{\beta}\left(1+\frac{(\beta+\delta)\mu}{\beta+1}+\frac{(\beta+\delta)(\beta+\delta-1)\mu^2}{(\beta+1)(\beta)\times 2!}+...\right)
\end{aligned}
$$

where $\tilde{F}$ is a hypergeometric function and $(.)_s$ is a Pochhammer symbol: a falling factorials. Substitute the hypergeometric representation of the incomplete Beta function, we can write the ratio in equation 11 as:

$$
\frac{B(\mu,\beta+1,\delta)}{B(\mu,\beta,\delta)} = \frac{\beta}{\beta+\delta}\left(1-\frac{1}{\tilde{F}(\beta+\delta,1;\beta+1;\mu)}\right)
$$

When δ is much larger than β, $\tilde{F}$ is very large. Consider that in our case, δ is the history of sales minus recall, and it is typically several times larger than β. Moreover, the *starting value* of δ ranges from about 5 million to 145 million across industries and $\delta_{js,t}$ grows larger in every period. We can safely say that $\tilde{F}$ will be negligibly small in any period of time, and the mean of standard Beta is a good proxy for the mean in truncated Beta distribution:

$$\mathbb{E}_t[\gamma] = \frac{B(\mu, \beta_t + 1, \delta_t)}{B(\mu, \beta_t, \delta_t)} \approx \frac{\beta_t}{\beta_t + \delta_t}.$$

The intuition of this approximation is that when δ is much larger than β, the Beta distribution is right-skewed with a thin right tail. When δ is large, the tail is very thin, and truncating on the right and a small upward shift of p.d.f. will have close to no effect on the expectation. Once we establish that truncated Beta is conjugate and its mean can be approximated with the standard Beta mean, the reputation updating process will just follow the derivation in section C.1.

C.3

$$\begin{aligned}
\mathbb{E}[u_{ijs,t}] &= \mathbb{E}\left[\mathbb{E}[u_{ijs,t}|x]\right] \\
&= \mathbb{E}\left[x(\log(I_{i,t} - p_{js,t}) + \alpha_j + \eta_{k,t} + \psi_{js,} + \epsilon_{ijs,t}) + (1-x)(\log(I_{i,t} - p_{js,t}) + \eta_{k,t} + \psi_{js,t} + \epsilon_{ijs,t})\right] \\
&= \mathbb{E}\left[\log(I_{i,t} - p_{js,t}) + \alpha_j x + \eta_{k,t} + \psi_{js,t} + \epsilon_{ijs,t}\right] \\
&= \log(I_{i,t} - p_{js,t}) + \alpha_j\, \mathbb{E}\left[x\right] + \eta_{k,t} + \psi_{js,t} + \epsilon_{ijs,t} \\
&= \log(I_{i,t} - p_{js,t}) + \alpha_j \int_{[0,1]} x\rho_{js,t}(x)dx + \eta_{k,t} + \psi_{js,t} + \epsilon_{ijs,t} \\
&= \log(I_{i,t} - p_{js,t}) + \alpha_j x_{js,t} + \eta_{k,t} + \psi_{js,t} + \epsilon_{ijs,t}
\end{aligned}$$

C.4 Market share prediction

Given that the idiosyncratic shock $\epsilon_{ijs,t}$ follows Type I extreme value distribution, the standard discrete choice model predicts the probability of consumer i choosing to buy from country k is:

$$\begin{aligned}
&\Pr(\mathbb{E}[u_{ijs,t}] > \mathbb{E}[u_{ijs',t}]|p_{js,t}, x_{js,t}, \eta_{k,t}, \xi_{js,t}) \\
&= \Pr\left(f(p_{js,t}, x_{js,t}, \eta_{k,t}, \psi_{js,t}; I_{ij,t}) + \epsilon_{ijs,t} > f(p_{js',t}, x_{js',t}, \eta_{k',t}, \psi_{js',t}; I_{ij,t}) + \epsilon_{ijs',t}\right) \\
&= \Pr(\epsilon_{ijs,t} - \epsilon_{ijs',t} > f(p_{js',t}, x_{js',t}, \eta_{k',t}, \psi_{js',t}; I_{ij,t}) - f(p_{js,t}, x_{js,t}, \eta_{k,t}, \psi_{js,t}; I_{ij,t}))
\end{aligned}$$

where $f(p_{js,t}, x_{js,t}, \eta_{k,t}, \psi_{js,t}; I_{i,t}) = \log(I_{i,t} - p_{js,t}) + \alpha_j x_{js,t} + \eta_{k,t} + \xi_{js,t}$ is the mean utility. Consumers are heterogeneous in term of expenditure on good j, and will only purchase one unit of good from country k if and only if $\mathbb{E}[u_{ijs,t}] > \mathbb{E}[u_{ijs',t}]$ for all $k' \neq k$.

C.5

Proof.

$$\begin{aligned}
&\mathbb{E}[\xi_{js,t}|p_{js,t}, x_{js,t}, z_{js,t}] \\
&= \mathbb{E}[\eta_{k,t} + \psi_{js,t}|p_{js,t}, x_{js,t}, z_{js,t}] \\
&= \mathbb{E}[\eta_{k,t}|p_{js,t}, x_{js,t}, z_{js,t}] + \mathbb{E}[\psi_{js,t}|p_{js,t}, x_{js,t}, z_{js,t}]
\end{aligned}$$

By law of iterative expectation:

$$\begin{aligned}
&\mathbb{E}[\psi_{js,t}|p_{js,t}, x_{js,t}, z_{js,t}] \\
&= \mathbb{E}\left[\psi_{js,t}|p_{js,t}, x_{js,t}, z_{js,t}, \eta_{k,t}\right]] \\
&= \mathbb{E}[0] = 0 \qquad (12)
\end{aligned}$$

And by the orthogonality between the instrument $z_{js,t}$ and $\eta_{k,t}$,

$$\mathbb{E}\left[\eta_{k,t}|p_{js,t}, x_{js,t}, z_{js,t}\right] = 0 \qquad (13)$$

Combine equation 12 and 13, we have

$$\mathbb{E}[\xi_{js,t}|p_{js,t}, x_{js,t}, z_{js,t}] = 0$$

□

C.6 Proof of Theorem 1

The following assumption is necessary for this proof. Intuition of this assumption is described in the main body and this is a formal layout.

Assumption 1 (Boundedness). *The parameters μ, β_0, δ_0 and realization of import flow $\{q_{js,t}\}_{t=1}^{T}$ and recalls $\{r_{js,t}\}_{t=1}^{T}$ satisfy the following:*

1. *$\mu \in [\underline{\mu}, 1]$, for some $\underline{\mu} > 0$. That is, the probability of recall given a bad product is bounded below by a positive number;*
2. *$\beta_0 > \underline{\beta_0} > 0$ and $\delta_0 > \underline{\delta_0} > 0$ for some $\underline{\beta_0}$, $\underline{\delta_0}$;*
3. *The quantity of import for each exporter and each product $q_{js,t}$ is nonnegative and bounded above by $\bar{q}_{js}$;*
4. *The units of products recalled $r_{js,t}$ do not exceed the units imported into the market in this period. Thus $r_{js,t}$ is nonnegative and bounded above by $\bar{q}_{js}$.*

Assumption 1 places almost no restrictions on the values of parameters in addition to those implied by the model intuition. Assumption 1-1 and 1-2 require that the parameters cannot take value zero. Lower bounds for the nonnegative parameters μ, β_0, δ_0 can be small, and its value will not affect my results. Assumption 1-3 and 1-4 specify that the data must be bounded above. Given that import flow depends on the exporters' production constraints and the importing country's wealth, there is no reason to believe that the volume of trade can be unlimited.

Assumption 2. *Let $\mathcal{H}_{jst}$ be the history when forming expectation for θ_{js} in period t. $\mathcal{H}_{jst} = \{(q_{js,t-1}, r_{js,t-1}), ..., (q_{js,0}, r_{js,0})\}$. The expectation for the quantity of product s from country j in the next period satisfies:*

$$\mathbb{E}\left[q_{js,t+1} | \theta_{js}, \mu, \mathcal{H}_{jst}\right] = \tilde{q}_{js}$$

That is, condition on history $\mathcal{H}_{jst}$, fraction of safe products θ_{js} and probability of recall for unsafe products μ, the expectation of import in period $t+1$ is time-invariant. Consumers do not learn about the size of market from history.

Assumption 2 is weaker than it seems. It states that consumers cannot predict the units of import in the following period from the history; but allows consumers to hold a belief that, say, China will in expectation sell more in next period than Cambodia. Consumers do not learn about the level of sales over time. There might be concerns that an exporter with superior production technology can produce both more reliable products (θ_{js} large) and at cheaper price. Those exporters will sell more. But this does not violate assumption 2 as long as the *expectation* of that advantage does not change over time. With assumption 1 and 2, we can conclude that learning is effective:

Proof. Let $\mathcal{H}_t$ be the history defined as $\mathcal{H}_t$ be the history when forming expectation for θ in period t. $\mathcal{H}_t = \{(q_{t-1}, r_{t-1}), ..., (q_0, r_0)\}$. The definition for x_t

$$x_t = \mathbb{E}\left[\theta \mid \mathcal{H}_t\right]$$

In this proof I drop all product and exporter subscript for cleanness of notation.

First, I will show that x_t is a martingale under assumption 1 and 2.

Conditional Expectation

Recall that $\gamma = \mu(1-\theta)$, thus given μ, if the sequence of conditional expectations of γ is a martingale, then the sequence of conditional expectations of μ is also a martingale. Define $\Gamma_t = \mathbb{E}[\gamma \mid \mathcal{H}_t]$ for simplicity. As shown in the proof in appendix C.1, the expectation of γ follows:

$$
\begin{aligned}
\mathbb{E}[\Gamma_{t+1} \mid \mathcal{H}_t] &= \mathbb{E}\left[\frac{\beta_t + \delta_t}{\beta_t + \delta_t + q_t}\Gamma_t + \frac{r_t}{\beta_t + \delta_t + q_t} \mid \mathcal{H}_t\right] \\
&= \mathbb{E}\left[\Gamma_t - \frac{q_t \Gamma_t}{\beta_t + \delta_t + q_t} + \frac{r_t}{\beta_t + \delta_t + q_t} \mid \mathcal{H}_t\right] \\
&= \Gamma_t + \mathbb{E}\left[\frac{q_t \Gamma_t - r_t}{\mu(\beta_t + \delta_t + q_t)} \mid \mathcal{H}_t\right] \\
&= \Gamma_t + \mathbb{E}\left[\mathbb{E}\left[\frac{q_t \Gamma_t - r_t}{\beta_t + \delta_t + q_t} \mid \mathcal{H}_t, q_t\right] \mid \mathcal{H}_t\right] \\
&= \Gamma_t + \mathbb{E}\left[\frac{1}{\beta_t + \delta_t + q_t}\mathbb{E}\left[q_t \Gamma_t - r_t \mid \mathcal{H}_t, q_t\right] \mid \mathcal{H}_t\right] \\
&= \Gamma_t + \mathbb{E}\left[\frac{1}{\beta_t + \delta_t + q_t}\left(q_t \Gamma_t - \mathbb{E}\left[r_t \mid \mathcal{H}_t, q_t\right]\right) \mid \mathcal{H}_t\right] \\
&= \Gamma_t + \mathbb{E}\left[\frac{q_t}{\beta_t + \delta_t + q_t}\left(\Gamma_t - \gamma\right) \mid \mathcal{H}_t\right] \\
&= \Gamma_t + \Gamma_t \, \mathbb{E}\left[\frac{q_t}{\beta_t + \delta_t + q_t} \mid \mathcal{H}_t\right] - \mathbb{E}\left[\frac{q_t \gamma}{\beta_t + \delta_t + q_t} \mid \mathcal{H}_t\right] \\
&= \Gamma_t + \underbrace{\left(\Gamma_t - \mathbb{E}\left[\gamma \mid \mathcal{H}_t\right]\right)}_{=0} \mathbb{E}\left[\frac{q_t}{\beta_t + \delta_t + q_t} \mid \mathcal{H}_t\right] \\
&= \Gamma_t \qquad (14)
\end{aligned}
$$

We can easily generalized the result in equation 14 to the case of $\mathbb{E}[\Gamma_{t+1} \mid \mathcal{H}_s], \quad s < t$. Note that $\mathcal{H}_0 \subset \mathcal{H}_1 \subset ... \subset \mathcal{H}_{t-1} \subset \mathcal{H}_t$.

$$
\begin{aligned}
\mathbb{E}[\Gamma_{t+1} \mid \mathcal{H}_s] &= \mathbb{E}\left[(\Gamma_{t+1} - \Gamma_t) + (\Gamma_t - \Gamma_{t-1}) + ... + (\Gamma_{s+1} - \Gamma_s) + \Gamma_s \mid \mathcal{H}_s\right] \\
&= \Gamma_s + \mathbb{E}\left[\underbrace{\mathbb{E}\left[\Gamma_{t+1} - \Gamma_t \mid \mathcal{H}_t\right]}_{=0} + \underbrace{\mathbb{E}\left[\Gamma_t - \Gamma_{t-1} \mid \mathcal{H}_{t-1}\right]}_{=0} + ... + \Gamma_{s+1} - \Gamma_s \mid \mathcal{H}_s\right] \\
&= \Gamma_s + \mathbb{E}\left[\Gamma_{s+1} - \Gamma_s \mid \mathcal{H}_s\right] \\
&= \Gamma_s \qquad \forall s < t \qquad (15)
\end{aligned}
$$

Since $\gamma = \mu(1 - \theta)$,

$$
\begin{aligned}
\mathbb{E}[\Gamma_{t+1} \mid \mathcal{H}_s] = \Gamma_s \quad \forall s < t &\Leftrightarrow \mathbb{E}[\mu(1 - x_{t+1}) \mid \mathcal{H}_s] = \mu(1 - x_s) \quad \forall s < t \\
&\Leftrightarrow \mathbb{E}[x_{t+1} \mid \mathcal{H}_s] = x_s \quad \forall s < t \qquad (16)
\end{aligned}
$$

Bounded

Next, I'd show that x_t is bounded given assumption 1: $\mu > \underline{\mu} > 0$.

The definition of the updating process guaranteed that x_t is bounded above. We can show this by way of induction. In the initial period, given $\mu > 0, \beta_0 > 0, \delta_0 > 0$,

$$x_1 = 1 - \frac{\beta_0}{\mu(\beta_0 + \delta_0)} < 1$$

For any period t, if $x_t < 1$, we have:

$$\begin{aligned}
x_{t+1} &= \frac{\beta_t + \delta_t}{\beta_t + \delta_t + q_t} x_t + \frac{\mu q_t - r_t}{\mu(\beta_t + \delta_t + q_t)} \\
&= \frac{(\beta_t + \delta_t)x_t + q_t - r_t/\mu}{\beta_t + \delta_t + q_t} \\
&< \frac{(\beta_t + \delta_t)x_t + q_t}{\beta_t + \delta_t + q_t} \\
&< \frac{\beta_t + \delta_t + q_t}{\beta_t + \delta_t + q_t} \\
&= 1
\end{aligned}$$

Thus x_t is bounded above by 1.

Given the positive lower bound for μ, β_0, δ_0 and upper bound for $\bar{q}$, we have:

$$\begin{aligned}
x_{t+1} &= \frac{\beta_t + \delta_t}{\beta_t + \delta_t + q_t} x_t + \frac{\mu q_t - r_t}{\mu(\beta_t + \delta_t + q_t)} \\
&> -\frac{r_t}{\mu(\beta_t + \delta_t + q_t)} \\
&> -\frac{r_t}{\underline{\mu}(\beta_t + \delta_t + q_t)} \\
&= -\frac{r_t}{\underline{\mu}(\beta_0 + \delta_0 + \sum_{\tau=1}^{\tau=t} q_\tau)} \\
&> -\frac{\bar{q}}{\underline{\mu}(\underline{\beta_0} + \underline{\delta_0})}
\end{aligned}$$

Thus x_t is bounded below by $-\dfrac{\bar{q}}{\underline{\mu}(\underline{\beta_0} + \underline{\delta_0})}$.

Boundedness implies that $\mathbb{E}[| x_t |] < \infty$, thus by definition 24.1 in Jacod and Protter (Jacod, Protter 2004), $\{x_t\}_{t=1}^T$ is a martingale. In addition, since x_t is bounded, we can conclude that it is also a uniformly integrable collection of random variables (see Definition 27.1 in (Jacod, Protter 2004)).

Now we have established that $\{x_t\}_{t=1}^T$ is a martingale and a uniformly integrable collection of random variables, then we can apply the martingale convergence theorem (see Theorem 27.3 in (Jacod, Protter 2004)) and conclude that

$$\lim_{t\to\infty} x_t = x_\infty \quad \text{exists a.s.}$$

Thus far, we have proved that $\{x_t\}_{t=1}^T$ converges and its limit exists almost surely. Next, I will show that the limit is indeed θ, the true fraction of bad products consumers are looking for.

Limit

Again, it may be easier to look at the limit of Γ_t first. Note that the existence of Γ_∞ can be proved using martingale convergence theorem as well, since we have shown that Γ_t is a martingale and Γ_t is bounded by 0 and 1.

$\square$