

PRODUCT RECALLS, IMPERFECT INFORMATION, AND SPILLOVER EFFECTS: LESSONS FROM THE CONSUMER RESPONSE TO THE 2007 TOY RECALLS

Seth Freedman, Melissa Kearney, and Mara Lederman*

Abstract—In 2007, the Consumer Product Safety Commission issued 212 recalls of toys and other children’s products, a sizable increase from previous years. We investigate changes in toy sales following these recalls. We find that for manufacturers that had recalls, unit sales of the types of toys involved in the recall fell relative to their sales of toys in other categories. We do not find evidence of within-manufacturer spillovers to dissimilar toys. We do, however, find large industry-wide spillovers in the form of sales losses to manufacturers that did not experience any recalls. Our results shed light on how consumers may draw inferences from information about product safety.

I. Introduction

IN markets with asymmetric information, the provision of information about product attributes can affect consumer demand and, in turn, firms’ incentives to invest in product quality. A growing empirical literature seeks to document the effects of information provision in markets in which consumers cannot verify product quality. Much of this literature focuses on quality disclosure and certification programs. These programs provide systematic information about quality for a large fraction of products in a given market (see Dranove & Jin, 2010, for a recent review). However, in many markets, including much of the consumer product market, consumers do not have access to this type of systematic information and instead are informed only when a particular product’s quality falls below a given threshold, say, through a product safety recall. Such information, while clearly more limited, may nevertheless affect market outcomes if consumers use this information to update their expectations about the quality of other products in the market.

This paper empirically investigates whether and how information about the quality of one set of products causes consumers to update their expectations about the quality of others. To do so, we estimate the consumer demand response to a set of highly publicized government-issued product recalls. The recall mechanism is the primary policy

tool used to regulate the safety of most consumer products. Recalls inform consumers of specific products that are in violation of a safety standard or pose a significant safety risk. However, recalls do not provide consumers with information about the safety of products available in the market since any products that are actively selling when they are recalled are immediately removed from retailers’ shelves. Moreover, in many cases, recalls are issued for products that are no longer active in the marketplace. As a result, to the extent that there is a consumer response to a recall, this would indicate that consumers use the information contained in the recall announcement to update their expectations of the safety of other products in the market. The goal of our empirical analysis is to document the level at which any such spillover effects are observed.

The specific context we examine is the series of highly publicized toy recalls that took place in 2007. In 2007, the Consumer Product Safety Commission (CPSC) issued 212 recalls of toys and other children’s products, as compared to 152 such recalls in 2006, 171 in 2005, and 121 in 2004.¹ This represents a greater than 80% increase in the number of recalled children’s items from 2006 to 2007 and a much larger increase than that is observed in other categories over this period. In surveys and interviews conducted at the time of the recalls, consumers clearly indicated an intention to change their behavior in response to these recalls. For example, in a Harris Poll of 2,565 adults in the United States conducted in October 2007, 33% of respondents said that they would buy fewer toys during the 2007 holiday season due to recent safety recalls and 45% said they would avoid toys from China (Harris Interactive, 2007). This wave of recalls and the resulting public concern about toy safety ultimately resulted in the passage of new federal legislation in August 2008: the Consumer Product Safety Improvement Act.

Several features of this setting make an examination of consumer expectations and spillover effects particularly interesting. First, the majority of the 2007 toy recalls involved risks associated with a common industry practice of producing in China, and many related specifically to the use of paint with high concentrations of lead. This feature raises the possibility that consumers took these announcements as information about the safety of an industry-wide practice rather than as information about the safety of any particular manufacturer’s toys. Second, licensing and branding are extremely common in the toy industry, with licensed products accounting for approximately one-quarter of toys in the industry (Clark, 2007). Trademarked charac-

Received for publication March 29, 2010. Revision accepted for publication September 23, 2010.

* Freedman: University of Michigan and Indiana University; Kearney: University of Maryland and NBER; Lederman: Rotman School of Management, University of Toronto

We gratefully acknowledge the helpful comments of Severin Borenstein, Jonathan Guryan, Judy Hellerstein, Ginger Jin, Arik Levinson, Soohyung Lee, Nuno Limao, and Abigail Wozniak, as well as seminar participants at the Energy Institute at the University of California (UC) Berkeley, University of Delaware, Georgetown, University of Michigan, Purdue, Syracuse, UC-Davis, UC-Irvine, and UC-Santa Barbara. We thank Danny Kim at NPD for answering our questions about the toy sales data and Kevin Mak at the Rotman Finance Lab for his assistance with assembling the stock price data. Molly Reckson provided capable research assistance. Financial support from the AIC Institute for Corporate Citizenship at the Rotman School of Management is gratefully acknowledged.

¹ Viscusi (1984) provides an institutional overview of the CPSC.

ters (such as Dora the Explorer) are often shared across different types of toys, as well as across toys produced by different manufacturers. Not only does this create another level at which consumers may draw inferences, but it also raises the possibility that imperfect information about manufacturer identify may prevent consumers from accurately acting on the inferences that they draw.² While we will not be able to test among competing explanations for the patterns we observe, we discuss the implications of each for both policy formulation and firm strategy. Finally, according to the CPSC announcements, 78% of the toys that were recalled in 2007 were not actively selling at the time of their recall. Thus, for most of these recalls, to the extent that they impose any costs on a manufacturer in the form of lost sales, these sales must be coming from other products.

Using the most comprehensive data available for this industry, we document how these recalls affected toy sales in the months following the recall announcements. Our analysis uses data on monthly infant/preschool toy sales from January 2005 to December 2007, inclusive. This broad category of toys is the largest in the industry and experienced the most recalls in 2007. Our empirical approach attempts to account for several important institutional features of the toy industry. In particular, the fact that toy sales are highly seasonal means that any demand response to a recall that is issued at any point in the year is most likely to occur at Christmas. However, the fact that the popularity or even life span of any particular toy may be short-lived means that what is popular one Christmas may not be popular the following year. It is thus very difficult to establish a single appropriate counterfactual level of sales that is clearly superior to alternative counterfactual estimates. We therefore carry out several complementary analyses that, together describe the patterns in the data. We begin by documenting differences in total toy sales across our three years of data. Then we investigate the relationship between having a recall during 2007 and Christmas 2007 sales. We follow standard industry practice and classify toys into categories (groupings of similar toys) and properties (groupings of toys that share a common brand or trademark) and estimate the impact of recalls on sales at the manufacturer-category level as well as the property-category level. Finally, we conclude by carrying out in-depth examinations of the largest and most widely publicized recalls from 2007.

Several key findings emerge from our analysis. First, the recalls had a negative impact on the Christmas season sales of the types of toys involved in the recalls. Our regression results indicate that relative to their categories that did not experience recalls, manufacturers' unit sales in categories that did have recalls were lower by about 30%. In addition, in the three recalls that we investigate in detail, the manu-

facturer's Christmas season sales in the affected category-property fell substantially. Thus, consumers appear to have used the information contained in the recall announcements to update their expectations about the safety of similar toys produced by the manufacturer involved in the recall. This suggests that a process of standards and recalls, such as that which regulates toy safety, can impose costs on firms in the form of reduced demand. These costs will provide at least some incentive for firms to invest in product safety.

Second, we find no evidence that a manufacturer's recall of one type of toy had any negative impact on its sales of other types of toys, beyond that experienced by nonrecalling manufacturers. This suggests that either consumers did not draw inferences from a manufacturer's recall of one type of toy about the safety of unrelated toys produced by that manufacturer or that, perhaps due to the prevalence of licensing and branding, they did not know which toys were produced by which manufacturer. If the latter is true, then the current process of recalls may need to be supplemented with additional information provision that enables consumers to better identify which toys are produced by which manufacturer. Of course, manufacturers may have incentives to limit association between their brands and publicize any recalls that do occur under a particular brand rather than the manufacturer name.³

Third, unit sales of infant/preschool toys produced by manufacturers that did not experience any recalls were about 30% lower at Christmas 2007 than in 2005, with no measurable change in 2006.⁴ Thus, the recalls appear to have had negative spillovers to the industry as a whole. Consistent with consumers' claims in surveys and in the media, this suggests that the specific recalls that took place led consumers to draw inferences about the overall safety of toys in the market.

Finally, our examination of the three highest-profile recalls of 2007 finds that recalls of branded toys may have positive or negative effects on the demand for other toys that share the brand. While our data do not allow a formal test of this, we speculate that the degree of similarity between the recalled toys and other toys in the property may affect the direction of the response.

³ Since we are measuring equilibrium changes in sales, it is also possible that diversified toy manufacturers took steps to shift consumer demand to their brands or product lines that were not involved in recalls. If this is the case, then the extent to which recalls provide large firms with incentives to invest in safety will depend on whether these firms have to undertake costly investments to prevent losses in unaffected categories.

⁴ Total spending on infant/preschool toys was down about 7% in 2007 as compared to 2006, with much of this decrease taking place in the third and fourth quarters. While we do not have quarterly sales data for any other industries, we can at least contrast the year-over-year spending change in toys to year-over-year changes in other industries. NPD Group press releases indicate that total spending on books was flat between 2006 and 2007, spending on apparel increased by about 3%, and spending on children's apparel increased by 6%. This suggests the reduction in sales of infant/preschool toys does not simply reflect reduced spending on consumer items. We provide additional information below on other changes in spending patterns based on data from the Consumer Expenditure Survey.

² For example, following a recall by Mattel, consumers may infer that all Mattel toys are less safe. However, if they do not know that toys produced under the Fisher-Price brand are in fact produced by Mattel, they will not be able to accurately act on that inference.

This paper is closely related to an existing, though mostly twenty-year-old, literature that measures the stock market response to recalls. The stock market response reflects the total costs that recalls impose on firms.⁵ Much of this literature focuses on drug and automobile recalls due to the high frequency of recalls in these industries. This literature includes Jarrell and Peltzman (1985), Pruitt and Peterson (1986), Hoffer, Pruitt, and Reilly (1988), Dranove and Olsen (1994), Barber and Darrough (1996), and Chu, Lin, and Prather (2005).⁶ With the exception of Hoffer et al. (1988), all of the papers find statistically significant negative stock price reactions to the recalls. Several of the papers compare the estimated drop in shareholder wealth to estimates of the direct costs of the recalls and find that the former exceeds the latter. They speculate that this excess loss is due to a loss of goodwill; this provides indirect evidence that the consumer response to recalls may be significant. Crafton, Hoffer, and Reilly (1981) and Reilly and Hoffer (1983) directly measure the demand response to automobile recalls.⁷ In a recent paper examining consumer demand responses, Cawley and Rizzo (2008) find that prescription drug withdrawals tend to have negative spillover effects to the utilization of other drugs within the same therapeutic class.

Our paper is also related to a more recent empirical literature that investigates the effects of government-mandated information disclosure programs.⁸ Economists have studied the impact of information disclosure policies on consumer and firm behavior in a variety of contexts, including restaurant hygiene grade cards (Jin & Leslie, 2003), nutritional labeling requirements (Mathios, 2000), mercury and fish consumption advisories (Shimshack, Ward, & Beatty 2007), SEC financial disclosure requirements (Greenstone, Oyer, Vissing-Jorensen, 2006), and environmental safety contexts (Benbear & Olmstead, 2008). To be clear, these papers measure the market response to information about available products' own quality. In contrast, we examine the impact of information about the quality of a particular set of products—typically products that are no

longer available for purchase—on consumer demand for other, potentially related products that are actively still available in the market.

The remainder of this paper is organized as follows. Section II provides background information. Section III describes the data. In section IV, we carry out our empirical analysis. Section V presents some additional considerations. We conclude with a discussion of the implications of our findings for both policy formulation and firm strategy.

II. Institutional Background

A. Toy Industry Basics

In 2005, the U.S. toy industry generated \$21.3 billion in retail sales.⁹ At both the manufacturer and retailer levels, the industry is dominated by a small number of large firms. At the manufacturer level, Mattel and Hasbro together account for roughly 30% of the market.¹⁰ The remaining firms are considerably smaller, with the third largest firm accounting for less than 4% of the market and the tenth largest firm accounting for just over 1% of the market.

For analysis purposes, the toy industry is classified into eleven supercategories, which are broad groupings of toys with similar uses or purposes. Examples include “action figures and accessories,” “infant/preschool,” and “youth electronics.” Supercategories are further subdivided into finer categories. The infant/preschool supercategory, which we focus on, is the largest in the industry, accounting for slightly more than 14% of total industry sales in 2005 (about \$3.2 billion). It is divided into thirteen finer categories such as “preschool vehicles” and “infant plush.” The four largest firms in the infant/preschool market are Mattel, Leapfrog, Hasbro, and RC2, which account for 27.3%, 10.3%, 6.6%, and 5.5%, respectively. The remainder of the infant/preschool market is served by firms accounting for small percentages.

Branding and licensing are quite common in the toy industry. A “property” refers to a set of toys that share a common brand. The property includes all toys produced by the owner of the brand, as well as all toys produced by firms that have licensed the rights to use the brand. There are two types of properties. The first type encompasses a brand that is owned by a toy manufacturer and used on some set of that manufacturer's toys. The manufacturer may license that brand to other toy manufacturers—but often does not—and may license that brand to firms producing other types of consumer products (for example, bicycles or children's furni-

⁵ Jarrell and Peltzman (1985) describe the multiple ways in which a product recall can lead to capital market losses: the direct costs of the recall in terms of inventory losses and refunds, costs of potential litigation, costs of changes in practices to improve quality and repair consumer goodwill, and lost profits due to decreases in consumer demand. All of these can provide incentives for firms to make costly investments in product safety.

⁶ A closely related literature estimates the stock market response to airline crashes. This literature includes Chalk (1987), Chance and Ferris (1987), Mitchell and Maloney (1989), Borenstein and Zimmerman (1988), and Bosch, Eckard, and Singal (1998). Borenstein and Zimmerman (1988) also investigate the impact of crashes on demand and find little or no effect.

⁷ They find that following a recall, there is a reduction in demand for the model type subject to the recall as well as for similar-sized cars produced by other manufacturers. They do not find evidence of negative demand spillovers to other cars produced by the manufacturer experiencing the recall.

⁸ There is a separate literature on quality certification, both voluntary and mandatory. This mechanism is used to alleviate informational deficiencies in such contexts as educational facilities and child care facilities. Hotz and Xiao (forthcoming) and Xiao (2010) are recent examples from this literature.

⁹ For comparison, in 2005, the U.S. book industry generated \$34.59 billion in sales, while the apparel industry generated \$181 billion. Video game hardware and software, which are not included within the definition of the “traditional” toy industry, generated \$10.5 billion in retail sales (Clark, 2007).

¹⁰ Though our study does not focus on retailers, we point out that in 2004, Walmart accounted for over 30% of toy sales by the top 25 retailers. Together, Walmart, Toys R Us, and Target accounted for over 60% of the sales by the top 25 retailers (Playthings, 2006).

TABLE 1.—RECALL TRENDS, 2004–2007

A: Number of Recalls by Consumer Product Category					
Year	Children's Products	Household Products	Outdoor Products	Sports and Recreation Products	Toys
2004	42	121	32	50	30
2005	64	122	28	76	31
2006	56	121	47	58	38
2007	130	132	38	64	82
B: Characteristics of Toy Recalls					
Year	Made in China	Choking	Laceration/Puncture	Lead	Magnets
2004	73.33%	50.00%	16.67%	13.33%	0.00%
2005	86.67	50.00	13.33	16.67	0.00
2006	75.68	48.65	8.11	10.81	5.41
2007	95.06	19.75	4.94	51.85	12.35

Source: Authors' tabulation from CPSC Web site.

ture). Mattel's Laugh & Learn brand is an example of this type of property. Mattel's Fisher-Price division produces approximately twenty different infant toys under the Laugh & Learn brand. The second type of property encompasses a brand that is owned by a firm outside the toy industry and is licensed to one or more toy manufacturers. In this case, the property would include all toys use the licensed brand or trademark and may include products from several different manufacturers. Examples include Spiderman, owned by Marvel Entertainment, and Dora the Explorer, owned by Nickelodeon. In some cases, a single toy manufacturer may obtain the exclusive rights to a license; in other cases, it may be shared by several different manufacturers.

B. The Recall Process

The recall process is initiated through one of three channels: a complaint made to the CPSC, a complaint made to the company whose product is in question, or a field sample or investigation.¹¹ When the CPSC receives a consumer complaint or is notified of a complaint made to a manufacturer, it immediately launches an investigation. If the content of the complaint is confirmed, the agency sends a letter to the company initiating a recall process. Manufacturers, importers, distributors, and retailers are required to report to the CPSC under section 15(b) of the Consumer Product Safety Act (CPSA) "within 24 hours of obtaining information which reasonably supports the conclusion that a product does not comply with a safety rule issued under the CPSA, or contains a defect which could create a substantial risk of injury to the public or presents an unreasonable risk of serious injury or death, 15 U.S.C. sec. 2064(b)."¹²

¹¹ This is based heavily on a description provided to us by a representative of the CPSC in a telephone conversation in April 2008.

¹² All of the CPSC toy recalls that we examined are listed on the CPSC Web site as "voluntary." Recalls that proceed along the channels described above are considered voluntary. A nonvoluntary recall would mean that the agency has to go through the legal system. The CPSC agent we spoke with could think of no such example in the past year of toy recalls. Furthermore, she could think of no instance in which a manufacturer initiated a recall of a product for which a violation had not been confirmed.

The large increase in the number of recalled toys and children's products observed in 2007, as compared to earlier years, is unique to this category of products. Panel A of table 1 reports the number of recalls per year in major categories of consumer products from 2004 through 2007. The number of toy recalls was 30, 31, and 38, respectively, for 2004, 2005, and 2006. That number jumped to 82 in 2007. For children's products, the numbers are 42, 64, and 56, with a jump up to 130 in 2007. The other categories do not show such a discrete increase in 2007.

III. Data

A. Recall Data

We collect details about the toy recalls that took place between 2004 and 2007 from the CPSC Web site. For each recall, the CPSC Web site lists the date of the recall; the product name; the number of units recalled; the importer, manufacturer, and/or distributor; a description of the hazard; details about any reported incidents or injuries; a description of the product to assist in identifying recalled items; details about where and when the item had been sold; the typical price; where the item was manufactured; and a picture of the item.¹³

Table 1, panel B shows changes in the characteristics of toy recalls from 2004 to 2007. There are two interesting patterns to note. First, although the majority of recalls in each year involve toys made in China, there is a noticeable increase in 2007, when 95% of recalls involved toys manufactured in China. Second, there is a change in the types of safety hazards leading to recalls. Prior to 2007, 13% of recalls were due to lead paint and 49% were due to choking; in 2007, these numbers were 52% and 20%, respectively.¹⁴

¹³ An appendix table in the NBER working paper version of this paper (Freedman, Kearney, & Lederman, 2009) lists all of the 2007 recalls and their major features.

¹⁴ One important difference between a lead hazard and a choking hazard is that a choking hazard may be discovered through the normal use of the toy, while a lead hazard will be discovered only through testing since the effects of lead on children are observed later in life.

The concern about lead paint in children's toys is largely driven by the fact that young children often put toys in their mouths and are thereby exposed to the lead content of paint. Lead is a powerful neurotoxin that interferes with the development of the brain and central nervous system, as well as the kidney and blood-forming organs. Lead poisoning in children is generally associated with behavioral problems, learning disabilities, hearing problems, and growth retardation.¹⁵ The federal legislation enacted in 2008 requires that surface lead, as in paint, must be below 90 parts per million as of August 2009, compared to the previous statutory level of 600 parts per million.¹⁶

B. Sales Data

We combine the recall data with data on the U.S. sales of toys in the infant/preschool toy supercategory from January 2005 through December 2007. We purchased these data from the NPD Group, self-described as the "single source for toy market research in the U.S., Europe, and Australia." The NPD data are based on a panel of more than 3 million consumers.¹⁷ The panel comprises two sets of consumers: an online panel of consumers who are instructed to record all their purchases, and a panel of consumers who have scanners in their homes and are supposed to scan everything they buy. From these two panels, NPD generates a toy-level data set with both actual data from the panels (for example, the number of transactions observed for each toy each month and the average price paid) and projected monthly unit and dollar sales figures (for the country). It is the latter measures that we use in our empirical analysis. After dropping observations for which no manufacturer information is available, our data set includes data from 156,524 transactions and 10,847 unique items over the full period.

Our data have three important features. First, the data are generally not reliable at the item level. Because the data are based on a sample of consumer purchases and because the toy industry is highly fragmented at the product level, most of the toys in the data set are involved in only a small number of transactions (NPD cautions against drawing inferences from cells with fewer than 35 transactions). In fact, the majority of toys have no transactions in any given month (84% of our toy-month cells have no recorded transactions). Because NPD does not keep records of market exit, we are

unable to determine whether zero transactions indicates that no consumers chose to purchase that toy in a given month or if the toy was no longer supplied.¹⁸ In the toy industry, new toys are introduced frequently, and current toys are either replaced or updated with new features, so exit may be an important consideration. For these reasons, we have no choice but to aggregate the item-level data over time or groups of items. In particular, we focus our regression analysis on sales at the level of the manufacturer-category and manufacturer-property. In our case studies, our level of analysis is generally the manufacturer-property category.

Second, our data do not include consumer-level variables. Therefore, though it would be interesting to explore consumer responses by retailer type or consumer demographics, we are unable to do so in this paper. Third, toy sales are highly seasonal. Roughly half of toy sales occur in the form of Christmas season purchases. An event-study type of methodology is thus inappropriate for analyzing these data because the demand response to a recall will likely not occur immediately. We instead focus our analysis on fourth-quarter sales, which include purchases made in October, November, and December of a given year.

IV. How Did Consumers Respond to the Toy Recalls?

Given the highly seasonal nature of toy sales, the time-varying popularity of particular brands, and our stated interest in estimating spillover effects, it is difficult to establish a single appropriate counterfactual for the sales that a given manufacturer, category, or property would have experienced at Christmas in 2007 absent the wave of recalls. We therefore carry out several complementary empirical analyses that together describe the patterns in the sales data in relation to the recalls. We begin by examining whether there are differences in aggregate toy sales across the three years of our sample using the NPD data. We compare what we find to trends in consumer purchases reported in the Consumer Expenditure Survey. We then run a series of regressions that relate sales changes at the level of the manufacturer-category (and property-category) to various measures of recall activity. We conclude with an in-depth analysis of the most highly publicized toy recalls from 2007.

Most of our analyses focus on changes in the equilibrium number of toys purchased rather than changes in equilibrium spending on toys because this allows us to directly investigate whether consumers purchased fewer toys (or fewer toys of a particular type) in Christmas 2007, as consumer surveys and the media predicted. Changes in dollar spending would also capture substitution to toys at different price points (for example, if consumers shifted to more expensive toys, perhaps out of a perception of greater

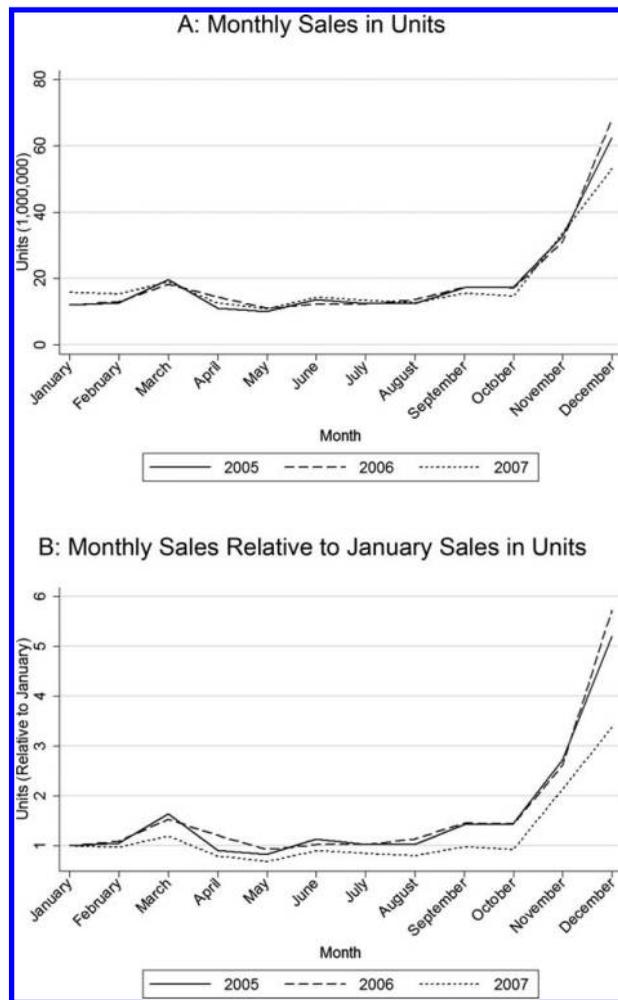
¹⁵ Information on the health risks of lead exposure for children are detailed on the National Safety Commission's Web page, <http://www.nsc.org/resources/issues/lead.aspx>, most recently accessed on March 9, 2009.

¹⁶ The recalls in 2007 varied in the extent to which they violated this standard. For example, Mattel's lead-related recalls in August and September 2007 involved toys with lead levels that ranged from just over the 600 ppm limit to 110,000 ppms (almost 200 times the applicable limit).

¹⁷ The ideal type of data for this project would be point-of-sale data collected directly from retailers. However, those types of data are no longer collected for the U.S. toy industry. The NPD Group previously collected such data but reported to us that in 2001, Walmart and Toys R Us stopped participating in this data collection. The loss of these retailers essentially negated the usefulness of the data.

¹⁸ Note that if a toy is recalled due to a faulty design and pulled completely from store shelves, sales of that particular toy would drop to 0. But in many of the 2007 recalls, only a particular production batch of the toy was recalled for having contained elevated levels of lead. Other batches of the toy would continue to be available to consumers.

FIGURE 1.—SALES OF INFANT/PRESCHOOL TOYS, 2005–2007



Panel A plots aggregate monthly infant/preschool toy sales in units for all manufacturers. Panel B scales units in each month by January sales in the respective year.

safety). For the sake of completeness, we do carry out some analysis with sales measured in dollars and describe any interesting differences that we find.

A. Changes in Total Toy Sales

We begin by exploring whether there are differences in monthly toy sales across the three years of data. Figure 1A plots total infant/preschool toy sales by month, for each of the three years in our sample. The figure indicates that Christmas sales in 2007 were lower than Christmas sales in each of the previous years. Specifically, unit sales in December 2007 were 22% lower than sales in December 2006 and 17% lower than sales in December 2005. September and October sales in 2007 also appear to be lower than sales in the same months in previous years, perhaps indicating some response to the recalls around the time of several of the largest, most widely publicized recalls (by RC2 and Mattel). Interestingly, sales at the beginning of 2007 were higher than in the corresponding months in 2005 and 2006,

with January 2007 sales being 33% greater than January 2006 sales. This is consistent with what we observe below when we look at the stock market performance of publicly traded toy firms: that toy firms outperformed the market for the first half of 2007.

These two observations suggest that absent the recalls, Christmas 2007 sales may have exceeded their 2006 levels and that a simple comparison of Christmas-to-Christmas sales changes could understate the actual impact on the industry. To examine this, in figure 1B, we plot monthly sales in each year normalized by January sales of that year. As is clear from the picture, once Christmas sales are normalized by January sales of the corresponding year, the gap between December 2007 and December 2006 becomes much larger. In 2006, December sales were 5.7 times January sales; in 2007, December sales were only 3.4 times January sales. While these figures show plots of only the raw data, they do suggest that the recalls may have had sizable effects on the infant/preschool category as a whole.

B. Changes in Other Categories of Spending

One might be concerned that reduced toy purchases during Christmas 2007 might merely reflect broader adjustments in consumer purchasing behavior during this time. Perhaps consumers just bought fewer items and gifts that season in general due to, say, macroeconomic conditions. To see whether the depressed sales that we see in the NPD data for infant/preschool toy sales during Christmas season 2007 extended to other consumer products, we examine data on household-level expenditures from the Consumer Expenditure Survey (CEX). We use data from the 2006 and 2007 Interview Survey files to examine household expenditures on children's items during Christmas season 2007 and compare this to fourth-quarter expenditures in 2006 and 2005 (the 2006 CEX file contains information about spending in the fourth quarter of 2005).¹⁹ We consider four broad categories of expenditures from the CEX detailed expenditure files. The first includes "toys, games, arts and crafts, trikes, and battery powered riders." Note that because expenditure amounts are not separately recorded for toys, as distinct from games, arts and crafts, riders, and so on, we cannot use these data to look specifically at infant/preschool toys or even all toys. The other categories of consumer goods we consider are (a) clothing: children and infant, (b) reading, and (c) TV, video hardware and software, and related equipment. Unfortunately the survey data do not separately record expenditures on children's reading material or children's TV and video products.

¹⁹ The CEX is conducted annually by the U.S. Bureau of Labor Statistics. The CEX interview survey is designed to collect data on major items of expense, household characteristics, and income. Each consumer unit in the sample is interviewed about its previous quarter's expenditures, reported by month, over a twelve-month period. We keep as our analysis sample for a given quarter households that have a complete calendar quarter (three months) of expenditures recorded.

TABLE 2.—QUARTERLY HOUSEHOLD EXPENDITURES ON SELECT CATEGORIES OF GOODS: CONSUMER EXPENDITURE SURVEY (CEX), INTERVIEW SURVEY DATA

	All Consumer Units (CUs)			CUs with Child(ren) below 18 Years			CUs with child(ren) below 18 Years, and Household Head a College Graduate		
	2005q4	2006q4	2007q4	2005q4	2006q4	2007q4	2005q4	2006q4	2007q4
Toys, games, etc.	54.0 (155.3)	57.1 (163.3)	72.6 (183.9)	110.6 (222.9)	105.0 (219.3)	139.3 (246.6)	123.5 (205.6)	121.9 (224.3)	167.9 (247.8)
Clothing: Children and infants	87.0 (306.0)	75.5 (177.6)	74.1 (167.8)	194.7 (475.2)	164.3 (235.9)	158.2 (209.8)	280.5 (794.6)	197.0 (270.1)	187.2 (228.4)
Reading	35.5 (71.4)	31.7 (62.3)	32.8 (68.9)	35.7 (74.7)	28.9 (59.6)	28.9 (63.0)	69.5 (111.2)	56.8 (87.3)	53.8 (92.5)
TV video hardware and software, and related equipment	273.1 (501.5)	262.1 (402.4)	275.8 (470.7)	344.8 (465.3)	317.8 (430.1)	344.0 (519.2)	437.0 (550.3)	410.1 (513.9)	436.5 (663.2)
Sample size	2,571	5,579	5,403	871	2,020	1,883	254	557	572

This table reports mean quarterly expenditures based on authors' tabulations of data from the Bureau of Labor Statistics (BLS) Consumer Expenditure Survey (CEX) Interview Survey files from 2006 and 2007, accessed through ICPSR. Data are aggregated from monthly expenditure data at the UCC level on the MTAB files. The "toys, games, etc." category is defined to be the UCC code for "toys, games, arts and crafts, trikes, and battery-powered riders." Monthly expenditures are CPI adjusted to 2007 dollars using the BLS All Urban Consumers, All-Items series.

Table 2 tabulates mean quarterly expenditures on these four select categories of spending, reported in 2007 dollars for all years. As shown in the table, the data do not provide any evidence of a reduction in 2007 quarter 4 expenditures on these items as compared to 2006 quarter 4 expenditures. The mean expenditure amounts in absolute terms increase for all but the clothing category, and in that case, the actual mean amount drops insignificantly from \$75.5 to \$74.1. In fact, the numbers in the table indicate that consumers spent more on the category "toys, games, etc." in the fourth quarter of 2007 than they did in the fourth quarter of 2006. Given our finding of reduced infant/preschool toys sales, this increase in expenditures in this category could be explained by consumers' shifting purchases toward other items in this broad category, such as games or more expensive items such as tricycles and battery-powered riders.

To push a bit on the possibility of substitution toward more expensive children's items, one might conjecture that a shift to more expensive children's products would be more likely among more educated (or higher-income) households. To consider this, we look separately at households headed by a college graduate versus those headed by an individual with a high school diploma or less. Among households with children that are headed by a noncollege graduate (not reported), mean quarterly expenditures in the "toys, games, etc." category increased 27%; among households with children headed by a college graduate, there was a 38% increase, from \$122 to \$168. Thus, we do observe a greater increase in spending among more highly educated households.

The findings suggest that consumers reduced their purchases of toys during the fourth quarter of 2007, consistent with anecdotal and poll reports of what consumers anticipated doing that Christmas season. Data on household expenditures give some indication that this reduction in infant/preschool toy purchases did not merely reflect

reduced spending on consumer goods during this quarter. We next consider the link between sales of infant/preschool toys and the recalls of 2007.

C. The Relationship between Recalls and Christmas Season Sales

Changes in sales at the manufacturer level. We turn now to an investigation of the relationship between recall activity in 2007 and subsequent toy sales. We start by looking at manufacturer-level unit sales. Table 3 reports changes in Christmas season sales for the total infant/preschool toy market (column 1), the top ten firms in our data (columns 2 through 11), and the two firms in our data that manufacture outside China and have a sufficient number of transactions (columns 12 and 13). We compare these changes in sales to the firm's recall activity in 2007.

Table 3 indicates that total unit sales in the infant/preschool toy market were down 13% in the fourth quarter of 2007 as compared to the fourth quarter of 2006 (recall from above that unit toy sales in December were down 22%). We also construct an adjusted-sales measure that scales fourth-quarter sales by first-quarter sales of the same year. This measure is appealing in that certain toys, manufacturers, or properties might have been on an upward trend and would have, in the absence of the wave of toy recalls, experienced higher sales in Christmas 2007 as compared to Christmas season 2006. For example, news stories report that RC2 was expecting a strong Christmas season in sales before its highly publicized recall in June 2007. We include this measure in table 3. For the infant/preschool category as a whole, adjusted fourth-quarter sales were lower by about 25%. When we define sales in terms of dollars, not reported in the table, the unadjusted change is -12.05%, and the adjusted change is -16.56%.

However, at the firm level, there is no apparent relationship between the number of toy recalls experienced and a

TABLE 3.—UNIT SALES (000s) FOR TOTAL MARKET, TOP TEN FIRMS, AND TOP TWO FIRMS MANUFACTURING OUTSIDE CHINA

	Top Ten Firms Based on 2005 Infant/Preschool Units Sold										Outside China		
	Total Market	Mattel	Leapfrog	Hasbro	RC2	Vtech	MGA Ent	Poof Toy Products	Tomy	Jakks Pacific	Russ Berrie	American Plastic Toys	Playmobil
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
2006 q1 units sold	43,176	10,727	4,091	2,342	2,824	1,465	1,081	548	1,089	553	864	577	372
2006 q4 units sold	116,356	32,579	11,495	9,888	6,441	7,194	3,670	2,280	2,594	2,448	1,337	578	986
2007 q1 units sold	49,937	10,819	4,827	3,037	4,867	2,017	1,267	1,021	1,710	1,225	1,019	468	369
2007 q4 units sold	101,640	27,014	8,999	9,632	5,707	7,272	3,319	2,384	1,050	3,378	889	732	717
% change	-12.65%	-17.08%	-21.71%	-2.59%	-11.40%	1.08%	-9.56%	4.56%	-59.52%	37.99%	-33.51%	26.64%	-27.28%
Adjusted % change													
2007 q4/q1 versus 2006 q4/q1	-24.47%	-17.79%	-33.65%	-24.88%	-48.59%	-26.58%	-22.84%	-43.88%	-74.22%	-37.71%	-43.62%	56.14%	-26.69%
Transactions 2006q4	22,826	6,623	2,370	1,924	1,411	1,403	701	561	532	465	259	68	175
2007 toy recalls		12	0	2	3	0	0	0	0	1	0	0	0

Based on authors' tabulations of NPD infant/preschool sales data on units sold. The total market column includes total infant/preschool sales based on NPD data.

firm's change in sales.²⁰ Mattel, by far the largest producer in the infant/preschool supercategory, had twelve recalls in 2007, yet its 2007 Christmas season unit sales decreased only 17% relative to Christmas season 2006 (its dollar sales, not reported, fell by 11%). In fact, of the four top 10 firms that had recalls—Mattel, Hasbro, RC2, and Jakks Pacific—only RC2 experienced a loss in sales larger than the general loss for the total market. Moreover, some of the largest unit sales losses were borne by firms that did not have any recalls of their own.

One possible explanation for this lack of a relationship between recalls and a named manufacturer's total level of infant/preschool sales is that firms are diversified across categories to varying degrees. So if consumers infer that a particular category of toys poses a safety risk, consumers may substitute from that category to other categories of toys. Firms that are highly concentrated in the affected category will experience large sales losses, while firms that are diversified across categories may experience nonlosses or even increases in sales in other categories. To be clear, recalls can have both a negative industry spillover effect, by which consumers reduce their purchases of all toys because of an updated expectation of risk, and a positive substitution effect, by which consumers substitute away from recalled items or categories to nonrecalled items or categories. Put differently, consumers may buy fewer toys altogether, but when they do buy, they shift their purchases to avoid toys or categories that have experienced recalls. If substitution happens at the level of the category and not the manufacturer, then manufacturers that are diversified across categories may actually experience smaller-than-average sales losses. Diversified manufacturers may also be able to encourage this substitution by offering lower prices or promotions in unaffected categories or by making fixed-cost investments in rebuilding their brand name.

We investigate this speculative explanation informally by documenting the diversification across categories among the top-producing firms. Table 4 reports the share of a manufacturer's 2006 fourth-quarter sales across the thirteen infant/preschool categories. The bottom row reports the calculated Herfindahl-Hirschman index (HHI), defined as the sum of the squares of shares over categories. The most diversified firms are Mattel (HHI of 1,863), Hasbro (HHI of 1,809), and MGA (HHI of 1,776). Playmobil produces only in the figures and playsets category, yielding an HHI of 10,000. Relevant to the patterns in table 3, as compared to Mattel, RC2 and Tomy are heavily concentrated in one category: preschool vehicles. A potential implication of this is that a shift away from purchases of preschool vehicles (say, following the recall of RC2 Thomas and Friends trains) would mean heavy sales losses for RC2 and Tomy, with no positive substitution into alternative categories to offset these losses.

²⁰ The CPSC recall announcements do not specify the category of the recalled toy. We therefore count all toy recalls in this exercise.

TABLE 4.—UNIT SHARE BY CATEGORY FOR MAJOR INFANT/PRESCHOOL TOY PRODUCERS

Category	Mattel	Leapfrog	Hasbro	RC2	Vtech	MGA Ent.	Poof Toy Products	Tomy	Jakks Pacific	Russ Berrie	American Plastic Toys	Playmobil
Other infant	27.0%	22.9%	17.5%	8.8%	20.8%	20.7%	0.0%	1.3%	2.8%	25.8%	0.0%	0.0%
Other PS	2.1	0.0	29.5	1.9	0.0	14.1	99.4	5.4	0.9	0.2	0.2	0.0
Bath	3.4	0.0	1.0	4.0	0.0	5.3	0.6	2.3	0.0	47.6	0.0	0.0
Infant plush	3.7	4.4	4.1	2.3	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0
Mobiles	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0
PS electronic learning	6.2	72.6	7.2	0.0	79.2	0.0	0.0	0.0	1.0	0.0	0.0	0.0
PS figures and playsets	28.9	0.0	12.4	0.1	0.0	2.1	0.0	0.0	0.0	0.0	0.0	100.0
PS Learning	1.0	0.1	2.0	1.2	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0
PS musical instruments	1.2	0.0	0.1	0.0	0.0	12.9	0.0	0.0	0.0	0.0	0.0	0.0
PS role playing	6.0	0.1	5.1	2.8	0.0	15.2	0.0	0.0	84.0	0.0	67.0	0.0
PS talking & sound	9.3	0.0	1.7	2.8	0.0	2.9	0.0	0.4	10.9	0.0	0.0	0.0
PS vehicles	10.2	0.0	19.3	72.2	0.0	26.6	0.0	90.5	0.4	0.0	32.9	0.0
Rattles/toy teethers	0.9	0.0	0.0	4.0	0.0	0.0	0.0	0.0	0.0	25.2	0.0	0.0
HHI	1,863	5,811	1,809	5,347	6,701	1,776	9,878	8,230	7,182	3,572	5,564	10,000

Authors' tabulations of NPD Infant/Preschool sales data on units sold in quarter 4 of 2006. PS = Preschool. HHI = Herfindahl-Hirschman index.

Manufacturer-category regressions. The data in tables 3 and 4 suggest that manufacturers involved in recalls in 2007 may have experienced different sales changes in their affected and unaffected categories and that, at least for some manufacturers, their unaffected categories may have actually performed better than manufacturers that did not have any recalls of their own. To explore these relationships more precisely, we estimate a series of manufacturer-category regressions. We again focus on Christmas season sales measured in units. All of our models include fixed effects for each manufacturer-category and dummy variables for the years 2006 and 2007. To estimate the effects of the 2007 recalls, we construct a measure of a manufacturer's overall level of recall activity in a given year ($Recall_{mt}$) and a measure of its recall activity in the particular category in a given year ($Recall_{mct}$) and interact these measures with the year 2007 dummy variable. Thus, at an intuitive level, our empirical approach is to estimate the change in sales in 2007 (relative to a manufacturer-category's mean level of sales) for three groups of manufacturer-categories: those by manufacturers that had no recalls at all, those by manufacturers that had recalls but in a different category, and those that had recalls in the specific manufacturer-category. While our regression specifications are similar in setup to that of a differences-in-differences model, we do not refer to it as such because our interest in spillovers means that we are explicitly assuming the absence of a control group. Rather, our model should be thought of as estimating differential changes in sales for manufacturer-categories with different levels of recall activity ("none," "in other categories," "in this category").

We include a manufacturer-category in the regression sample if the manufacturer has at least 35 transactions in the fourth quarter of a given year. We additionally require that the manufacturer-category has positive fourth-quarter sales in all three years.²¹ We identify a recall as belonging to one of the thirteen infant/preschool categories if the item in a CPSC recall announcement appears in our sales data. In other words, if a particular item does not appear in the NPD sales data, we make the assumption that it is outside one of these categories (likely because it is outside the infant/preschool supercategory). Note that such a recall would still be reflected in the indicator variable for a recall to the manufacturer. In our first set of regressions, we capture recall activity with simple indicators for whether the manufacturer-category experienced at least one recall during the calendar year and for whether the manufacturer experienced at least one recall in any category during the calendar year.²² We explore alternate measures of recall activity below.

²¹ Our regression sample includes 43 manufacturers and 203 manufacturer-categories. The average fourth-quarter sales in our sample is 451,565 units, or \$6,624,871.

²² We do not include an indicator for recalls at the category level because recalls occur in twelve of the thirteen categories in our data. The indicator would therefore be highly collinear with the 2007 year dummy.

We estimate the following regression equation:

$$\begin{aligned} \text{Log}(\text{Units}Q4_{mct}) &= \beta_0 + \beta_1(\text{Recall}_{mct} \times \text{yr}07_t) + \beta_2(\text{Recall}_{mt} \times \text{yr}07_t) \\ &+ \beta_3(\text{Recall}_{mct} \times \text{yr}06_t) + \beta_4(\text{Recall}_{mt} \times \text{yr}06_t) \\ &+ \beta_5(\text{Recall}_{mct} \times \text{yr}05_t) + \beta_6(\text{Recall}_{mt} \times \text{yr}05_t) \\ &+ \beta_7(\text{yr}07_t) + \beta_8(\text{yr}06_t) + \mu_{mc} + \varepsilon_{mct}. \end{aligned} \quad (1)$$

The coefficient β_7 measures the change in fourth-quarter sales for manufacturer-categories by manufacturers that had no recalls in 2007 and might reasonably be considered a measure of industry-level spillover effects. Of course, the causal interpretation of this coefficient is subject to the assumption that toy sales would otherwise have been flat relative to year 2005. While this assumption is a strong one, it is bolstered by the descriptive data presented above showing that sales of other consumer products did not show decreases in Christmas season 2007.

The coefficient β_2 measures the conditional change in fourth-quarter sales for unaffected categories by manufacturers that had at least one recall in 2007. It indicates whether unaffected categories of manufacturers that had recalls had a systematically different change in sales in 2007 than categories of manufacturers that did not have any recalls. This coefficient reflects the within-manufacturer spillover effect.

The other main coefficient of interest is β_1 , which measures the conditional change in fourth-quarter 2007 sales for manufacturer-categories that experienced a recall during 2007. It indicates whether the affected categories of manufacturers that had recalls experienced a different change in sales in 2007 than the unaffected categories of those manufacturers. This coefficient captures the most direct spillover effect of the product recalls, as these toys are presumably the most similar to those recalled.

A key assumption of our empirical approach is that absent the recalls, differences in quarter 4 sales across manufacturer-categories in 2007 are adequately captured by the manufacturer-category fixed effects. Put differently, the key identifying assumption of our model is that there are no unobserved factors that are correlated with the recalls and that lead to differential time trends across manufacturer-categories. To the extent that recalls are correlated with time-invariant (over three years) unobserved differences in manufacturer size or quality, these would be captured by the manufacturer-category fixed effects that we include. The estimates of the coefficient on the manufacturer-category recall variable could be a biased estimate of the effect of a recall if the likelihood of a recall in particular manufacturer-category is correlated with the growth rate of that category. However, if faster-growing manufacturer-categories are more likely to have recalls, then this would bias us against finding a negative effect of the recall on manufacturer-category sales.

Finally, although we believe that the 2007 toy recall context is quite distinct from toy recalls in earlier years (because the 2007 recalls were highly publicized and dominated by lead recalls related to toys made in China), we also include indicators for 2005 and 2006 recalls interacted with 2005 and 2006 year dummies. We do this for the purpose of econometrically controlling for potentially systematic, confounding movements in the data. However, we note that in these early years, we have considerably fewer recalls to use to identify these variables. Indeed, there are only two recalls in 2005 to manufacturer-categories included in our data. Moreover, while the effects of the 2007 recalls are identified relative to a manufacturer-category's sales in the two years previous to the recall, the effects of earlier recalls are identified relative to sales in at least one year subsequent to the recall. Thus, it is more difficult to assign a causal interpretation to these coefficients.

Main regression results. The first column of table 5 reports the results of equation (1). The estimated coefficient on the year 2007 indicator suggests that for manufacturers that had no recalls in 2007, unit sales in a category were, on average, down 27 log points (30.9%) compared to the fourth quarter of 2005. There is no measurable difference between 2006 and 2005. The point estimate on the indicator variable for having a recall at the level of manufacturer-category in 2007 is negative (32.8 log points, or 38.9%) while the point estimate on the indicator variable for having a recall at the level of the manufacturer is positive (32.6 log points or 38.5%).²³ Both are statistically significant at the 10% level. This pattern of estimates suggests that in 2007, manufacturers that experienced recalls had lower sales in their affected categories (relative to their unaffected categories) but higher sales in their unaffected categories (relative to manufacturers that had no recalls at all).²⁴

To confirm whether this pattern of sales changes is likely to have been caused by the recalls, we estimate whether the manufacturers and manufacturer-categories that experienced recalls in 2007 were following different sales patterns in 2006. That is, we look for evidence of "pretrends" in the affected manufacturers and manufacturer-categories by including an interaction between the 2007 manufacturer-category and manufacturer recall indicators with the year 2006 indicator variable. As reported in column 2, the coefficient on the interaction between the manufacturer-category recall indicator and the year 2006 dummy is -0.049 and statistically indistinguishable from 0. This implies that in 2006, sales in the manufacturer-categories that experienced

²³ We find similar results in the recall variables when we replace the year dummies with category-year fixed effects, which control for differential growth rates across categories.

²⁴ The estimated β_5 shows that manufacturer-categories that experienced a recall in 2005 saw sales increases in 2005 quarter 4 sales relative to their manufacturer-category fixed effect (which are calculated over subsequent years in this case). We do not think this is interpretable as a causal effect of the 2005 recalls.

TABLE 5.—OLS IMPACT OF A RECALL DURING THE YEAR ON MANUFACTURER-CATEGORY QUARTER 4 UNITS SOLD

Dependent Variable: Log(Units)	All Firms		Top Fifteen Firms
	(1)	(2)	(3)
I(07 Recall to Man/Cat)	-0.328*	-0.341	-0.385*
× I(2007)	(0.189)	(0.219)	(0.212)
I(07 Recall to Man/Cat)		-0.049	
× I(2006)		(0.212)	
I(07 Recall to Manuf)	0.326*	0.458**	0.432*
× I(2007)	(0.182)	(0.207)	(0.239)
I(07 Recall to Manuf)		0.330*	
× I(2006)		(0.199)	
I(06 Recall to Man/Cat)	-0.118	-0.070	0.073
× I(2006)	(0.281)	(0.275)	(0.239)
I(06 Recall to Manuf)	-0.058	-0.192	-0.004
× I(2006)	(0.160)	(0.166)	(0.224)
I(05 Recall to Man/Cat)	0.529**	0.484**	0.539**
× I(2005)	(0.202)	(0.231)	(0.117)
I(05 Recall to Manuf)	-0.257*	-0.201	-0.051
× I(2005)	(0.145)	(0.163)	(0.171)
I(2007)	-0.269**	-0.297**	-0.227
	(0.099)	(0.102)	(0.144)
I(2006)	-0.080	-0.124	-0.008
	(0.097)	(0.103)	(0.169)
Constant	11.447**	11.440**	12.296**
	(0.066)	(0.066)	(0.089)
N	609	609	258
R ²	0.879	0.880	0.914
Number of Manuf/ Categories		203	86
Number of I(07 Recall to Man/Cat)		13	11
Number of I(07 Recall to Manuf)		56	45
Number of I(06 Recall to Man/Cat)		6	5
Number of I(06 Recall to Manuf)		59	52
Number of I(05 Recall to Man/Cat)		2	1
Number of I(05 Recall to Manuf)		26	13

Firm ranks determined by total units sold in 2005. Robust standard errors in parentheses. Regressions include manufacturer-category fixed effects. Statistically significant at the *10%, **5%. Man/Cat=Manufacturer category. Manuf=Manufacturer.

recalls in 2007 were not statistically different from the manufacturers' other categories (after accounting for mean differences captured by the fixed effects). This suggests that the sales decrease that we measure for the affected manufacturer-categories for 2007 is likely to have been caused by the recalls.²⁵ This claim is further strengthened by looking at alternative measures of recall activity, as described below.

The coefficient on the interaction between the 2007 manufacturer recall indicator (not manufacturer-category indicator) and the 2006 dummy is positive and statistically

significant, indicating that in 2006, sales in nonrecalled categories among manufacturers that experienced recalls in 2007 were following different sales trends than manufacturers that had no recalls in 2007. A test of the hypothesis that the coefficients on the 2006 and 2007 interactions are identical cannot be rejected at the 5% level. We interpret this as indicating that in 2007, the unaffected categories of manufacturers that experienced recalls in 2007 appear to be have been following their 2006 trend. Thus, we see no evidence that consumers "punished" these manufacturers by reducing purchases of the manufacturer's items in unaffected categories. Put differently, we do not detect any within-manufacturer spillovers. This may be because consumers do not change their expectations of the safety of toys in unaffected categories (beyond what they inferred for all toys) or because consumers are simply not aware of which toys are produced by which manufacturer.

In the final column of table 5, we report the results for our main specification estimated on a subsample of the top fifteen manufacturers in our data, based on 2005 unit sales. The results are quite similar, though the coefficient on the year 2007 indicator variable is less precisely estimated.²⁶ This is perhaps not surprising since only four of the fifteen firms in this subsample have no recalls in 2007. In results not reported, we run the same regressions with dollars as the dependent variable instead of units. Coefficient estimates are qualitatively similar.

We have also estimated a series of property-category regressions, analogous to those described above. These models allow us to estimate the effects of recalls that involved a toy that was part of a property. We identify recalls that are part of properties if the CPSC recall announcement mentions the property's name in its description of recalled toys. For the sake of space, we do not present those analyses in detail.²⁷ In brief, the point estimates on the recall indicators in the property-category level regressions follow a similar pattern to those reported in table 5, but they are not statistically significant. Specifically, the point estimate on the indicator variable for having a recall in the property-category is negative (-0.131, standard error of 0.249) and the point estimate for having any recall in the property is positive (0.066, standard error of 0.212). This would be consistent with consumers shifting away from a particular type of toy (in the category) in the property after such a toy was recalled but not shifting away from dissimilar toys within the property. However, the data lack the precision to detect within-property spillovers.

²⁵ One might question whether the negative coefficient on the manufacturer-category recall indicator might be a mechanical result due to the fact that the recalled toys are removed from the market. This is not a substantive concern in our sample because we are looking only at fourth-quarter sales and only one of the recalls that appear in our data involved a toy that was actively selling in the fourth quarter of 2007. We have estimated our main specification without this recall and obtain an estimate of β_1 of -0.293 with a standard error of 0.190.

²⁶ The magnitude and precision of the coefficient on the year 2007 indicator variable also vary with the criteria used for inclusion of a manufacturer-category in the regression sample. When we require manufacturer-categories to have a larger number of transactions in the NPD data, our sample size falls, and, not surprisingly, the precision of the estimates is reduced. However, we always find a negative point estimate on the 2007 indicator. The point estimates range from -0.055 to -0.269.

²⁷ This set of results is available from the authors on request.

TABLE 6.—OLS IMPACT OF A RECALL DURING THE YEAR ON MANUFACTURER-CATEGORY QUARTER FOUR UNITS SOLD, BY LEVEL OF RECALL ACTIVITY AND LEVEL OF NEWS COVERAGE

<i>Dependent Variable: Log(Units)</i>	Recalled Units			News Coverage
	(1)	(2)		(3)
I(High Recall Activity Man/Cat)	-0.422*	-0.536**	I(High Coverage Man/Cat)	-0.492**
	(0.242)	(0.236)		(0.179)
I(High Recall Activity Man)	0.414*	0.443*	I(High Coverage Man)	0.337
	(0.244)	(0.250)		(0.210)
I(Low Recall Activity Man/Cat)	-0.330	-0.021	I(Medium Coverage Man/Cat)	-0.550**
	(0.227)	(0.326)		(0.273)
I(Low Recall Activity Man)	0.225	0.180	I(Medium Coverage Man)	0.500*
	(0.201)	(0.209)	I(Low Coverage Man/Cat)	0.333
				(0.282)
			I(Low Coverage Man)	0.127
				(0.213)
<i>N</i>	609	609	<i>N</i>	609
<i>R</i> ²	0.880	0.880	<i>R</i> ²	0.880
Number of I(Large Recall Man)	36	36	Number of I(High Coverage Man/Cat)	8
Number of I(Large Recall Man/Cat)	4	9	Number of I(High Coverage Man)	13
Number of I(Small Recall Man)	50	50	Number of I(Medium Coverage Man/Cat)	2
Number of I(Small Recall Man/Cat)	9	4	Number of I(Medium Coverage Man)	40
			Number of I(Low Coverage Man/Cat)	3
			Number of I(Low Coverage Man)	43

Level of recall activity is determined by whether a manufacturer (MAN) or manufacturer-category (Man/Cat) experienced recalls over the previous year totaling above or below 1 million units in column 1 and \$20 million in column 2. News coverage is measured by counting the number of articles mentioning the name of the company and the words *toy* and *recall* in the LexisNexis database of major United States and world publications during the thirty days following the recall. Article counts are summed over all recalls to a manufacturer or manufacturer-category during a given year. Low news coverage corresponds to recalls with 0 to 9 articles, medium coverage corresponds to recalls with 10 to 99 articles, and high coverage corresponds to recalls with 100 or more articles. Robust standard errors are in parentheses. Regressions include manufacturer-category fixed effects. Statistically significant at *10%, **5%.

Alternate measures of recall activity. So far we have measured recall activity with a simple indicator for whether a recall was experienced during the year. Recognizing that the set of recalls that took place during 2007 was heterogeneous in scale as well as resulting media attention and that consumers may react differently to larger or more widely publicized recalls, table 6 reports results from regressions using alternative measures of recall activity that give greater weight to larger or more salient recall events.²⁸ This set of regressions contributes to a consideration of whether the effects estimated in table 5 are interpretable as causally related to the recalls. In column 1, we include indicator variables for high- and low-recall activity during 2007, where high-recall activity is defined as a manufacturer-category (or manufacturer) having more than 1 million units recalled during 2007. In column 2, the level of recall activity is based on the dollar value of the recalls, and the high recall activity variables equal 1 for manufacturer-categories (or manufacturers) that recalled more than \$20 million worth of toys during 2007. In both specifications, the estimated pattern of coefficients is similar to what we find in table 5, and the precision of the estimates is improved. In addition, the point estimate on the indicator variables for a manufacturer-category having high recall activity is greater

in absolute value. The estimates on indicators for low recall activity are not significant.²⁹

Finally, we try to capture the salience of a firm's recalls with a measure of the amount of media attention that the recalls attracted. We measure news coverage using the LexisNexis database of Major U.S. and World Publications. We count the number of news articles mentioning the name of the company and the words *toy* and *recall* in the thirty days immediately following the recall. The article counts are summed over all recalls to a manufacturer or manufacturer-category in a given year. We use these sums to divide manufacturers and manufacturer-categories into three categories of media coverage. Coverage is categorized as low for manufacturer-categories (or manufacturers) with 0 to 9 news stories, medium for those with 10 to 99 news stories, and high for those with 100 or more news stories. The results using these measures of recall activity are presented in the third column of table 6. The coefficient on the indicator for a manufacturer-category having high media coverage is -0.492 and is highly significant. The coefficient on the medium coverage variable is also large and negative and statistically significant at the 10% level. The data do not indicate an effect on sales of low coverage recalls. We

²⁸ We categorize recall activity only in 2007 and simply control for a manufacturer-category or manufacturer having a recall in 2005 or 2006 as we do in table 5. The coefficients on the 2005 and 2006 recall variables, as well as the coefficients on the year dummies, are not reported in table 6 but are very similar to their values in table 5. These results are available from the authors on request.

²⁹ In the first column, the pattern of coefficients on the low-recall activity indicators is similar to the pattern on the high-recall-activity indicators, but this is not the case in the second column, where the coefficients are much smaller. This difference arises because the categorization based on units recalled classifies fewer manufacturer-categories as having high-recall activity than does the categorization based on dollars recalled. Thus, some manufacturer-categories with high-recall activity based on dollars are identifying the low-recall-activity indicator in the first specification.

interpret these results as indicating that there was a greater consumer response to those recalls that garnered significant media attention. We are cautious about attributing a causal role to media coverage in this setting since it may be endogenous; however, understanding the extent to which the media can and do influence consumer expectations of product quality is clearly an important avenue for future research.³⁰

Before turning to our detailed case studies, it is worth briefly commenting on the extent to which our results may be capturing demand or supply effects. To be clear, we are estimating the equilibrium change in sales from one Christmas to another. We are relating this change in sales to recalls that took place somewhere between Christmas 2006 and 2007, with most of the large recalls taking place between August and September 2007. While the sales changes we estimate could, in theory, capture movements in both supply and demand, the timing of manufacturers' decisions suggests that significant changes in supply between the occurrence of a recall and Christmas 2007 are unlikely. Industry sources indicated to us that at least for large manufacturers, development decisions for the products they will offer at a particular Christmas usually begin about one and half years in advance, so decisions about what toys to offer for Christmas 2007 would have begun in the spring of 2006. Retailers typically place orders about one year in advance, so Christmas 2007 orders would have been placed at Christmas 2006. Actual production of the toys (most of which takes place in China) begins several months later, which indicates production lags of at least several months in this industry. Given this, when a manufacturer has a recall in, say, August 2007, that manufacturer is limited in the extent to which it can affect supply for the upcoming Christmas. For example, it is unlikely to be able to switch subcontractors and have production ready in time for Christmas sales. In addition, given that more than 75% of our recalls involved toys that were no longer selling at the time of the recall, the incentive for manufacturers to make supply-side changes would appear to be weaker. Finally, to the extent that manufacturers reallocate advertising dollars from recalled categories of toys to other categories of toys, this should be considered a demand shifter, and our estimates should be interpreted as capturing the net effect of the recalls on demand accounting for any other changes manufacturers (or retailers) make to shift demand.

C. An In-Depth Examination of Three High-Profile Recalls

The results in table 6 suggest that the consumer response to these recalls was driven by the largest and most widely publicized recalls that took place. In this section, we complement our regression analyses with a detailed examina-

tion of the demand response to the three highest-profile recalls of 2007. Honing in on particular cases in this way has the added advantage of allowing us to consider effects at the level of manufacturer-category-property. Data limitations preclude us from using regression techniques to estimate with statistical precision impacts at this triple-interaction level.³¹

On June 13, 2007, and September 26, 2007, RC2 announced two separate recalls of Thomas the Tank Engine wooden trains, buildings, and other train set items. The first recall involved 1.5 million toys, and the second involved an additional 200,000 toys. Both were a result of excessive levels of lead paint. On August 2, 2007, Mattel recalled 967,000 various figures and other toys sold under the Fisher-Price brand because of excessive lead in the surface paint. Most of the toys involved in the recall were part of the Sesame Street and Dora the Explorer properties. These recalls received significant media attention, as noted in Appendix B. Furthermore, each of these recalls involved an extremely popular property. Thomas the Tank Engine is the second largest infant/preschool property, and Sesame Street and Dora the Explorer are, respectively, the seventh and eighth largest properties in the supercategory.

Table 7 considers the RC2 recalls of toy trains and accessories in the Thomas the Tank Engine railroad property. We start by examining what happened to RC2's sales in the affected category-property, Vehicles, produced under the Thomas & Friends brand. The first column of the table shows that RC2's adjusted Christmas season sales (fourth-quarter sales divided by first-quarter sales) of Thomas vehicles decreased by 58.5% in 2007 (unadjusted sales decreased by 11%). Consistent with our regression results, we observe a decrease in the sales of the firm's toys that are most similar to those that were recalled. RC2's non-Thomas vehicles experienced a similar sales decline (column 3), which suggests that consumers substituted away from the RC2 vehicle category, and the sales loss in this category is not property specific.³² In contrast, RC2's sales outside the affected category and affected property (column 4) increased slightly over this period. This is consistent with our findings that firms are not experiencing sales losses in unaffected categories and may even be experiencing sales increases.

³¹ Six properties are named in at least one lead recall in 2007: Thomas and Friends (RC2 recalls, June and September 2007); Dora the Explorer, Sesame Street, and Go Diego Go! (Mattel recall, August 2007); GeoTrax (Mattel recall, September 2007); and Baby Einstein (Kids II recall, October 2007). We focus on three of these. We do not include detailed examinations of the other three property recalls for the following reasons: The Go Diego Go! property did not exist in the first half of 2006, our Lexis-Nexis search finds only three articles mentioning the Kids II Baby Einstein recall, and the GeoTrax property is an exclusive Mattel brand and therefore is not produced by other manufacturers. In addition, the GeoTrax recall involved fewer than 100,000 units. Comparable case study tables for these events are available on request.

³² There are too few RC2 nonvehicles Thomas toys in the data to permit an examination of spillover effects within the manufacturer and property but outside the category.

³⁰ Shimshack et al. (2007) consider the role of education and newspaper readership on the consumer response to the FDA's advisor about the consumption of tuna.

TABLE 7.—RC2 RECALL OF THOMAS AND FRIENDS, JUNE 2007 AND SEPTEMBER 2007: QUANTITY CHANGES (000S OF UNITS)

	RC2				Non-RC2			
	Thomas & Friends		Non-Thomas & Friends		Thomas & Friends		Non-Thomas & Friends	
	Vehicles	Nonvehicles	Vehicles	Nonvehicles	Vehicles	Nonvehicles	Vehicles	Nonvehicles
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2006 q1 units sold	1,452	-	377	989	1,213	160	2,622	36,357
2006 q2 units sold	1,319	-	358	785	1,242	122	2,577	31,214
2006 q3 units sold	1,785	-	320	718	1,034	163	3,111	35,877
2006 q4 units sold	3,860	-	790	1,726	2,738	603	9,230	97,344
2007 q1 units sold	3,097	-	900	861	1,787	290	3,099	39,894
2007 q2 units sold	2,291	-	645	819	1,063	363	2,566	29,814
2007 q3 units sold	2,005	-	554	1,039	1,200	158	2,305	34,210
2007 q4 units sold	3,420	-	736	1,535	2,303	566	8,705	84,359
2007 q4 versus 2006 q4 units change	-440	-	-54	-191	-435	-37	-525	-12,985
2007 q4 versus 2006 q4 % change	-11.40%	-	-6.84%	-11.07%	-15.89%	-6.14%	-5.69%	-13.34%
Adjusted % change	-58.46%	-	-60.97%	2.16%	-42.91%	-48.21%	-20.20%	-21.02%
Transactions 2006q4	866	6	182	357	520	141	1,881	18,873

Lead-related recall of Various Thomas and Friends Wooden Railway Toys; 1.5 million units worth \$60 million (June recall) and 200,000 units worth \$5 million (September recall). Figures are not calculated for cells with fewer than 35 transactions.

Next, we examine how RC2's recalls affected sales to competitors' products within the property and within the category. Column 5 of the table indicates that adjusted Christmas season sales of Thomas vehicles produced by firms other than RC2 were down 42.9%. Sales of Thomas items outside the vehicles category were also down by more than 40%. Sales of items outside the manufacturer, category, and property (column 8) decreased by 21%, which is similar to our findings in our regressions. The data do not show a larger loss in sales for non-RC2/non-Thomas vehicles as compared to non-RC2/non-Thomas sales outside the category (column 7 versus column 8), which suggests consumers were not substituting away from toys in the category that were neither produced by RC2 nor produced under the Thomas brand. To summarize, in the case of RC2's recalls, in addition to the direct effect of the recall on the affected manufacturer-category-property, we also see negative effects on sales in the manufacturer-category (outside the property) and on the property (outside the manufacturer, within and outside the category).

This finding that consumers moved away from non-RC2 Thomas items at twice the rate of non-Thomas items suggests that either consumers used RC2's recalls to update their expectations about the safety of all Thomas toys or they were confused about which Thomas items were included in the recall. While we cannot formally test between these hypotheses, we point out that the RC2 recall is a case where consumer confusion could easily arise because the Thomas items produced by the various different manufacturers sharing the Thomas license are quite similar.³³

Table 8 conducts a similar exercise for Mattel's recall of Dora the Explorer items. The first column of the table shows that Mattel's adjusted sales of Dora figures and play-

sets decreased by 53%. This again provides evidence of a large direct effect of a recall on the affected manufacturer-category-property. As in the RC2 case, Mattel's adjusted sales in the category but outside the property also decreased, in this case by about 38% (column 3). Both of these numbers are substantially larger than the overall 17% sales decrease that Mattel experienced (from table 3). However, Mattel's adjusted sales outside the category and outside the property (column 4) fell by only 12%. Consistent with what we have found earlier, this again suggests that there is no net negative spillover to the manufacturer's sales outside the category and property.

Perhaps the most interesting patterns in table 8 appear in columns 2 and 6. These columns look at the change in sales of Dora items outside the figures and playsets category, so spillovers across categories within the property.³⁴ Column 2 indicates that Mattel's sales of Dora items in unaffected categories did not decrease and actually increased slightly. Column 6 indicates that rivals' sales of Dora items in unaffected categories increased by more than 40%.

These numbers suggest that after Mattel's recall of various Dora figures and playsets, consumers did not decrease their purchases of other Dora products, but instead substituted specifically toward other types of Dora toys. Furthermore, this suggests that consumers did not interpret Mattel's recall as providing information about the safety of all Dora items. Nor were they confused by Mattel's Dora recall; rather, they interpreted it as providing information about the safety of specific Dora items. These effects contrast with what we found in the case of RC2's recall where we observed that rivals' sales of Thomas items (within and outside the affected category) decreased. Note, however, that there is less heterogeneity in the types of toys produced

³³ Rashes (2001) provides evidence of a similar type of confusion on the part of investors. Specifically, he finds that investors get confused between stocks with very similar ticker symbols and mistakenly execute trades on one firm in response to information about the other.

³⁴ It appears that Mattel has exclusive licensing rights to produce figures and playsets in the Dora brand, as there are no sales of Dora figures and playsets made by other manufacturers.

TABLE 8.—MATTEL'S DORA RECALL (AUGUST 2007): QUANTITY CHANGES (000s OF UNITS)

	Mattel				Non-Mattel			
	Dora the Explorer		Non-Dora the Explorer		Dora the Explorer		Non-Dora the Explorer	
	Figures and Playsets	Nonfigures and Playsets	Figures and Playsets	Nonfigures and Playsets	Figures and Playsets	Nonfigures and Playsets	Figures and Playsets	Nonfigures and Playsets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2006 q1 units sold	872	431	2,430	6,994	-	225	1,217	31,007
2006 q2 units sold	856	288	1,951	5,413	-	178	732	28,201
2006 q3 units sold	425	280	2,459	7,565	-	150	791	31,339
2006 q4 units sold	1,246	1,071	8,163	22,099	-	820	3,217	79,740
2007 q1 units sold	486	470	2,576	7,287	-	216	1,365	37,537
2007 q2 units sold	243	357	1,298	6,400	-	146	575	28,542
2007 q3 units sold	344	301	1,511	6,849	-	317	800	31,353
2007 q4 units sold	329	1,192	5,336	20,157	-	1,113	1,898	71,615
2007q4 versus 2006q4 units change	-917	121	-2,827	-1,942	-	293	-1,319	-8,125
2007 q4 versus 2006 q4 % change	-73.60%	11.30%	-34.63%	-8.79%	-	35.73%	-41.00%	-10.19%
Adjusted % change	-52.62%	2.06%	-38.34%	-12.46%	-	41.39%	-47.40%	-25.81%
Transactions 2006q4	215	198	1,800	4,410	0	135	579	15,489

Lead-related recall of Sesame Street, Dora the Explorer, and other children's toys; 967,000 units worth \$21.8 million. Figures are not calculated for cells with fewer than 35 transactions.

TABLE 9.—MATTEL'S SESAME STREET RECALL (AUGUST 2007): QUANTITY CHANGES (000s OF UNITS)

	Mattel				Non-Mattel			
	Sesame Street		Non-Sesame Street		Sesame Street		Non-Sesame Street	
	Recalled Categories	Nonrecalled Categories						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
2006 q1 units sold	965	-	8,557	1,115	88	-	26,377	5,982
2006 q2 units sold	805	-	6,874	775	38	-	22,805	6,264
2006 q3 units sold	762	-	8,752	1,179	63	-	25,862	6,332
2006 q4 units sold	3,141	-	26,829	2,511	232	-	66,984	16,309
2007 q1 units sold	1,298	-	8,680	809	68	-	31,105	7,890
2007 q2 units sold	794	-	6,627	851	43	-	22,836	6,365
2007 q3 units sold	851	-	7,334	810	53	-	25,155	7,237
2007 q4 units sold	2,011	-	22,636	2,322	195	-	58,159	16,208
2007 q4 versus 2006 q4 units change	-1,130	-	-4,193	-189	-37	-	-8,825	-101
2007 q4 versus 2006 q4 % change	-35.98%	-	-15.63%	-7.53%	-15.95%	-	-13.17%	-0.62%
Adjusted % change	-52.40%	-	-16.82%	27.45%	8.77%	-	-26.37%	-24.65%
Transactions 2006q4	562	30	5,468	563	52	35	12,412	3,704

Lead-related recall of Sesame Street, Dora the Explorer, and other children's toys; 967,000 units worth \$21.8 million. Recalled categories include all other infant toys, bath toys, electronic learning, figures and playsets, learning toys, role playing, talking and sound, and vehicles. Figures are not calculated for cells with fewer than 35 transactions.

under the Thomas brand than under the Dora brand. Most Thomas items are trains or train-related accessories. In contrast, the Dora items that were not in the affected figures and playsets category were as diverse as umbrellas, a Dora kitchen, and Dora electronic learning toys. The patterns in this table are not consistent with broad confusion about recall details—remembering the brand but not the specific toy—but they are consistent with (arguably) reasonable inferences about product safety.

Finally, table 9 considers Mattel's recall of Mattel Sesame Street figures. As in table 8, the data indicate that sales of toys by Mattel in the affected category-property fell by 52.4%, roughly twice as much as the general decrease in toys and three times as much as Mattel's overall decrease. Mattel has exclusive licensing rights to much of the Sesame Street brand, so there is not much scope to explore broader

effects on the property. The data again fail to show any evidence of a net negative spillover to Mattel sales outside the category and property. To the contrary, there was a 27.5% increase in Mattel's adjusted infant/preschool toy sales outside the affected categories and property. This observation is consistent with consumers substituting from affected to unaffected categories in a way that favors more diversified firms. The last two columns of the table indicate sales decreases for other manufacturers outside the property that are quite similar to the 25% that we have found above.

In summary, these focused case studies show that (a) in all three cases, there was a large decrease in adjusted sales in the affected manufacturer-category-property; (b) there were negative spillovers to the manufacturer's sales within the category but no apparent negative spillover to the manufacturer's sales outside the category or property; and (c)

there were negative spillovers to rivals' sales in the affected property when the types of items included in the property are very similar (the Thomas case) but positive spillovers to rivals' sales in the property when the types of items inside the property are dissimilar.³⁵ These patterns are broadly consistent with consumers drawing reasonable inferences about toy safety.

V. Additional Considerations

A. The Stock Market Response to Toy Recalls

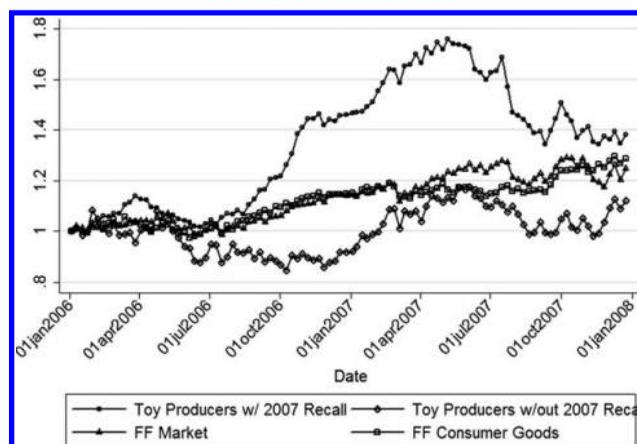
In this section, we briefly describe our investigation of the impact of the recalls on the stock market performance of publicly traded toy manufacturers. We use data on daily stock market prices from the Center for Research in Security Prices (CRSP) accessed through Wharton Research Data Services (WRDS). We obtain daily end-of-day stock quotes between 2004 and 2007 for each firm identified as a toy manufacturer and listed on any of the three major U.S. exchanges. To identify toy manufacturers, we use Mergent Online, a database of business characteristics, to collect primary and secondary Standard Industry Classification (SIC) codes for firms. We identify firms as toy manufacturers if any of their SIC codes fall in categories 3942 (Dolls and Stuffed Toys) or 3944 (Games, Toys, and Children's Vehicles, Except Dolls and Bicycles). We identify eighteen such firms. Many of the recalls named firms that are not publicly traded, so we cannot conduct our analysis on the full set of toy recalls.

Figure 2 plots trends in four stock market indices: toy producers with a 2007 recall, toy producers without a 2007 recall, Fama-French market index, and Fama-French consumer good index.³⁶ The data show that toy firms that experienced recalls outperformed relative to the market index until mid-2007 and then greatly underperformed relative to the market index, with toy firm stock prices falling drastically while the market showed no break in trend. Consistent with the findings from the sales data, the index of toy firms that did not experience recalls follows a similar pattern, but to a

³⁵ For all three case studies, the toys that were listed in the recall were still for sale at the time of the recall, though not necessarily the particular batch implicated in the recall. Although our case studies focus on fourth-quarter sales, we do present data from all four quarters, which means that we include the quarter of the recall. To account for any mechanical effect on sales that results from the recalled items being pulled from store shelves, we have recalculated changes in adjusted sales excluding items that are listed in the recall announcements and that appear in the sales data. Doing so generally reduces the effect on the manufacturer-category-property but results in the same qualitative conclusions. Note, however, that to the extent that a recall covers only a particular batch of toys, it will not result in all units of a toy being pulled from the shelves and our correction will overcompensate for any mechanical effect.

³⁶ The indices created by Fama and French include a market index of all NYSE, AMEX, and NASDAQ firms and an index of firms producing consumer goods. The other two are indices that we created for toy producers. They are value-weighted indices of the eighteen toy manufacturers traded on the three major U.S. exchanges in our data. We separate these into separate indices for the six firms that had at least one recall in 2007 and the twelve firms that did not. All indices are normalized to 1 on January 1, 2006, and are adjusted for stock splits and dividends.

FIGURE 2.—STOCK MARKET INDICES OF MAJOR TOY PRODUCERS COMPARED TO MARKET INDICES



Toy producers with and without 2007 recalls are value-weighted indices of the eighteen firms traded on the NYSE, AMEX, or NASDAQ that have primary or secondary SIC codes indicating toy production. Seven of these firms had a 2007 recall, and 11 did not. All other indices are from Kenneth French's Website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html (accessed February 13, 2009)). FF Market includes all firms on the three major U.S. exchanges. FF Consumer Goods include firms producing various consumer products.

lesser degree. In contrast, the stock market performance of consumer good firms moved very close to the market. This decline in market performance of toy firms over the third quarter of 2007 coincides with the increasing frequency of toy recalls and two other patterns of the recalls. First, eight of the ten recall events in the second half of 2007 were lead related, whereas only one had been prior to this period. Second, this period was characterized by much higher press coverage of recalls. There had been very few news articles covering earlier recalls, but the 2007 recalls received wide media coverage. For example, there were 551 articles in the thirty days following Mattel's August 14, 2007, recall.

We next conduct an event study analysis that allows us to identify if this decline can be linked to specific recall announcements. We conduct this event study in the spirit of Jarrell and Peltzman (1985) following the methodology laid out by MacKinlay (1997). Our event study sample includes 25 recalls that took place between 2004 and 2007. These recalls involved eight firms.

The details of our methodology, as well as detailed tables of results, are available in an online appendix. For the sake of space, we only briefly describe our main approach and results here. In brief, the basic strategy of an event study is to estimate the relationship between the affected firm's daily stock return and an index (or set of indices) of market performance over an estimation window, which is a period of time preceding the event. These parameters are then used to calculate the predicted returns to the affected firm during the event window, which is a period of time surrounding the event. Abnormal returns are calculated as the difference between the actual returns and the predicted returns over the event window. These represent the impact of the event, or the news, on the firm's market value. Abnormal returns can be calculated for each day in the event window, or they

can be summed over a given interval to obtain an estimate of the cumulative abnormal returns (CAR) for that event.

For the 2007 recalls, our event study analysis yields an estimated average abnormal return of -1.1 percentage points on the day of the recall, which is statistically significant at the 5% level. However, a closer look at the data reveals that this is entirely driven by one event, the September 2007 RC2 recall. Furthermore, the average two-day CAR that we estimate is not statistically significant. An examination of ten-day CARs provides further evidence that the loss in shareholder wealth in the toy industry over the second half of 2007 is not characterized by a series of negative abnormal returns to a particular firm following its own recall announcements.

We interpret our stock market analysis as indicating that the relative decline in the stock market performance of toy manufacturers over the second half of 2007 is most appropriately characterized as a gradual investor response to a perceived industry-wide problem. We cannot discern in this paper whether the investor response reflects expectations about general consumer demand for toys or expectations of higher costs of regulatory compliance for the industry as a whole.

B. “Made in China”

A final consideration is the possibility that there was a market response specifically targeted at toy firms producing in China. The Harris Poll found that 45% of respondents indicated they would avoid buying toys manufactured in China (Harris Interactive, 2007). As it turns out, almost all infant/preschool toys in the United States are manufactured in China. So what did consumers actually do when it came to making toy purchases?

We attempt a simple investigation of whether consumers shifted toy purchases to toys made outside China.³⁷ We look at our toy sales data from 2005 quarter 1 to 2007 quarter 4 to see if there is an increase in the share of purchased toys manufactured outside China. In the infant/preschool supercategory of toys, a handful of notable toy manufacturers produce outside China, including Playmobil 1-2-3 (Malta, Germany), Haba (Germany), PlanToys (Thailand), Siku (Germany), Vikingtoys (Thailand), and Geomag (Switzerland). Some manufacturers that produce mainly in China advertise specific items that are “Made in the U.S.” We had two research assistants explore the Web sites of the top fifty manufacturers producing toys in our NPD sample to identify toys that are noted as being produced outside China. Under the assumption that when not otherwise noted, a toy was manufactured in China, we calculate the share of infant/preschool toys (measured in units) manufac-

tured outside China. Over our sample period, this share ranges from 2.5% to 4.9%. In 2007, the share was actually at the lowest end of the range; however, in contrast to previous years, it was higher in the fourth quarter of 2007 than in the third. We also examine directly the share of toy sales to the two notable non-Chinese producers in our data: American Plastic Toys and Playmobil. We see no obvious upward ticks in their sales trends.

VI. Conclusion

This paper has provided an examination of the consumer response to the highly publicized wave of 2007 toy recalls. Our analysis of sales data for the infant/preschool supercategory of toys reveals several interesting patterns. First, there is evidence that consumers responded to this wave of recalls by substituting specifically away from a manufacturer’s category of toys involved in a recall. This indicates that consumers used the information contained in the recall announcements to update their expectations of the safety of the types of toys involved in the recalls. This finding is important because it speaks to whether the costs of information gathering in potentially confusing contexts are prohibitive to consumer action. We document that in the case of the highly publicized toy recalls in 2007, the costs of information gathering were not prohibitive to eliciting a consumer response. We also find that the response was greatest for the largest and most widely publicized recalls. Although our data are not rich enough to estimate whether this is because these recalls involved the largest number of items, the largest firms in the industry, popular properties, or significant media attention, our analysis suggests that the consumer response to a recall is likely to depend on a variety of these factors.

Second, we find no evidence that consumers specifically shifted away from other types of toys produced by manufacturers’ involved in a recall. This may be because a manufacturer’s recall of one type of toy did not change consumers’ expectations of the safety of that manufacturer’s other toys (relative to toys in general) or that manufacturer association in this industry is particularly weak. Alternatively, it may be that because large, diversified manufacturers took measures to bolster sales of unaffected toy categories. Precisely understanding the mechanism that prevented manufacture-wide above-average losses in sales is important to understanding the incentive effects inherent in the recall process. It will also speak to the likelihood of observing within-manufacturer spillovers in other settings. These are interesting questions for future research.

Further examination of the most highly publicized lead-related infant/preschool toy recalls of 2007 suggests that recalls involving toys that are part of a property can have positive or negative spillovers to sales of rivals’ toys in the same property. The existence of shared brands in this industry generates externalities between competing licensees as well as licensees and the licensor. Licensees will have sub-

³⁷ An analysis conducted by the Federal Reserve Board of San Francisco investigates what happened to imports from China (Candelaria & Hale, 2008). This work finds that imports in the toy industry actually increased relative to forecast levels. Of course, because this work is carried out in 2008, it cannot detect whether the industry will experience long-term changes away from Chinese imports.

optimal incentives to invest in protecting the brand. This may have implications for the types of protections built into the licensing contracts and perhaps also for the likelihood of observing exclusive versus nonexclusive licenses. It also gives licensors incentives to invest directly in toy testing, which they have begun doing.³⁸

The prevalence of licensing in this industry also increases the likelihood that consumers are imperfectly informed about manufacturer identity. If consumers are unable to distinguish between manufacturers and brands, then the current process of recalls may be improved by providing additional information that enables consumers to better identify which toys are produced by which firms. Manufacturers that produce under a brand involved in a recall by another firm clearly have an incentive to inform consumers that it was not their branded products that were recalled. Manufacturers that produce under multiple brands may have incentives to limit association between their brands and publicize any recalls that do occur under a particular brand rather than the manufacturer name.

Finally, we find that consumers reduced overall infant/preschool toy purchases in Christmas 2007. This is consistent with consumers' responding to this particular set of toy recalls by updating their expectations about the safety hazards of toys in general. Similarly, our examination of stock market data for this industry also reveals an industry-wide loss of value, relative to several different benchmarks. These findings indicate that even information about a specific set of products that is no longer available in the marketplace can have large effects on consumers' behavior. While we do not have data on the costs of improved safety, the finding that a relatively small number of recalls by a few large manufacturers appears to result in decreased sales and capital market losses for the segment as a whole suggests that from an industry perspective, investments in safety may be too low. When a shared industry practice is involved, such as production in China in the case of the toy recalls, the potential for spillover effects appears to be especially large.

³⁸ See, for example, <http://www.nytimes.com/2007/09/10/business/media/10toys.html>.

REFERENCES

- Barber, Brad M., and Masako N. Darrough, "Product Reliability and Firm Value: The Experience of American and Japanese Automakers, 1973–1992," *Journal of Political Economy* 104:5 (1996), 1084–1099.
- Benneer, Lori S., and Sheila M. Olmstead, "The Impacts of the 'Right to Know': Information Disclosure and the Violation of Drinking Water Standards," *Journal of Environmental Economics and Management* 56:2 (2008), 117–130.
- Borenstein, Severin, and Martin B. Zimmerman, "Market Incentives for Safe Commercial Airline Operation," *American Economic Review* 78:5 (1988), 913–935.
- Bosch, Jean-Claude, E. Woodrow Eckard, and Vijay Singal, "The Competitive Impact of Air Crashes: Stock Market Evidence," *Journal of Law and Economics* 41:2 (1998), 503–519.
- Candelaria, Christopher, and Galina Hale, "Did Large Recalls of Chinese Consumer Goods Lower U.S. Imports from China?" *FRBSF Economic Letter*, no. 2008–17 (2008), 1–3.
- Cawley, John, and John A. Rizzo, "Spillover Effects of Prescription Drug Withdrawals," *Advances in Health Economics and Health Services Research* 19 (2008), 119–144.
- Chalk, Andrew, "Market Forces and Commercial Aircraft Safety," *Journal of Industrial Economics* 36 (1987), 61–81.
- Chance, Don M., and Ferris, Stephen P., "The Effect of Aviation Disasters on the Air Transport Industry," *Journal of Transport Economics and Policy* 21 (1987), 151–165.
- Chu, Ting-Heng, Che-Chun Lin, and Larry J. Prather, "An Extension of Security Price Reactions around Product Recall Announcements," *Quarterly Journal of Business and Economics* 44 (2005), 33–48.
- Clark, Eric, *The Real Toy Story* (New York: Free Press, 2007).
- Consumer Product Safety Act, Pub. L. No. 92–573 (1972).
- Crafton, Steven M., George E. Hoffer, and Robert J. Reilly, "Testing the Impact of Recalls on the Demand for Automobiles," *Economic Inquiry* 19 (1981), 694–703.
- Dranove, David, and Ginger Zhe Jin, "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature* 48 (2010), 935–963.
- Dranove, David, and Chris Olsen, "The Economic Side Effects of Dangerous Drug Announcements," *Journal of Law and Economics* 37:2 (1994), 323–348.
- Freedman, Seth, Melissa S. Kearney, and Mara Lederman, "Product Recalls, Imperfect Information, and Spillover Effects: Lessons from the Consumer Response to the 2007 Toy Recalls," NBER working paper no. 15183 (2009).
- Greenstone, Michael, Paul Oyer, and Annette Vissing-Jorensen, "Mandated Disclosure, Stock Returns and the 1964 Securities Act," *Quarterly Journal of Economics* 121:2 (2006), 399–460.
- Harris Interactive, "Recent Toy Recalls Threaten Sales of Chinese Products This Holiday Season," *Harris Poll* 114, November 14, 2007.
- Hoffer, George E., Stephen W. Pruitt, and Robert J. Reilly, "The Impact of Product Recalls on the Wealth of Sellers: A Reexamination," *Journal of Political Economy* 96:3 (1988), 663–760.
- Hotz, V. J., & M. Xiao, "Strategic Information Disclosure: the Case of Multi-Attribute Products with Heterogeneous Consumers," *Economic Inquiry* (forthcoming).
- Jarrell, Gregg, and Sam Peltzman, "The Impact of Product Recalls on the Wealth of Sellers," *Journal of Political Economy* 93:3 (1985), 512–536.
- Jin, Ginger Z., and Phillip Leslie, "The Effects of Information on Product Quality: Evidence from Restaurant Hygiene Grade Cards," *Quarterly Journal of Economics* 118:2 (2003), 409–51.
- Mathios, Alan D., "The Impact of Mandatory Disclosure Laws on Product Choices: An Analysis of the Salad Dressing Market," *Journal of Law and Economics* 43:2 (2000), 651–677.
- Mitchell, Mark L., and Michael T. Maloney, "Crisis in the Cockpit? The Role of Market Forces in Promoting Air Travel Safety," *Journal of Law and Economics* 32:2 (1989), 329–355.
- MacKinlay, A. Craig, "Event Studies in Economics and Finance," *Journal of Economic Literature* 35:1 (1997), 13–39.
- Playthings, "Top Toy Retailers, 2004," in R. S. Lazich (Ed.), *Market Share Reporter 2007*, Vol. 2 (Detroit, MI: Thomson Gale, 2006).
- Pruitt, Stephen W., and David R. Peterson, "Security Price Reactions around Product Recall Announcements," *Journal of Financial Research* 9:2 (1986), 113–122.
- Rashes, Michael S., "Massively Confused Investors Making Conspicuously Ignorant Choices," *Journal of Finance* 56:5 (2001), 1911–1927.
- Reilly, Robert J., and George E. Hoffer, "Will Retarding the Information Flow on Automobile Recalls Affect Consumer Demand?" *Economic Inquiry* 21:3 (1983), 444–447.
- Shimshack Jay P., Michael B. Ward, and Timothy K. M. Beatty, "Mercury Advisories: Information, Education and Fish Consumption," *Journal of Environmental Economics and Management* 53:2 (2007), 158–179.
- Viscusi, W. Kip, *Regulating Consumer Product Safety* (Washington, DC: AEI Press, 1984).
- Xiao, Mo, "Is Quality Certification Effective? Evidence from the Child-care Market," *International Journal of Industrial Organization* 28 (2010), 708–721.

This article has been cited by:

1. Shu He, Huaxia Rui, Andrew B. Whinston. 2018. Social Media Strategies in Product-Harm Crises. *Information Systems Research* 29:2, 362-380. [[Crossref](#)]
2. Sebastian Utz. 2017. Corporate scandals and the reliability of ESG assessments: evidence from an international sample. *Review of Managerial Science* 5. . [[Crossref](#)]
3. Kathleen Cleeren, Marnik G. Dekimpe, Harald J. van Heerde. 2017. Marketing research on product-harm crises: a review, managerial implications, and an agenda for future research. *Journal of the Academy of Marketing Science* 45:5, 593-615. [[Crossref](#)]
4. Ram Bala, Pradeep Bhardwaj, Pradeep K. Chintagunta. 2017. Pharmaceutical Product Recalls: Category Effects and Competitor Response. *Marketing Science* . [[Crossref](#)]
5. Anandasivam Gopal, Brad N. Greenwood. 2017. Traders, guns, and money: The effects of mass shootings on stock prices of firearm manufacturers in the U.S. *PLOS ONE* 12:5, e0177720. [[Crossref](#)]
6. Adam R. Fremeth, Guy L. F. Holburn, Brian K. Richter. 2016. Bridging Qualitative and Quantitative Methods in Organizational Research: Applications of Synthetic Control Methodology in the U.S. Automobile Industry. *Organization Science* 27:2, 462-482. [[Crossref](#)]
7. Veronica F. Pozo, Ted C. Schroeder. 2016. Evaluating the costs of meat and poultry recalls to food firms using stock returns. *Food Policy* 59, 66-77. [[Crossref](#)]
8. Julia Cagé, Dorothée Rouzet. 2015. Improving “national brands”: Reputation for quality and export promotion strategies. *Journal of International Economics* 95:2, 274-290. [[Crossref](#)]
9. Muhammad Saad Memon, Young Hae Lee, Sonia Irshad Mari. 2015. Analysis of Traceability Optimization and Shareholder's Profit for Efficient Supply Chain Operation under Product Recall Crisis. *Mathematical Problems in Engineering* 2015, 1-8. [[Crossref](#)]
10. ÁRON TÓTH. 2014. REPUTATION EFFECTS IN THE MARKET OF CERTIFIERS: EVIDENCE FROM THE AUDIT INDUSTRY. *Economic Inquiry* 52:2, 505-517. [[Crossref](#)]
11. Thomas J. Kniesner, John D. Leeth. Regulating Occupational and Product Risks 493-600. [[Crossref](#)]
12. J. Michael Collins, Kosali I. Simon, Sharon Tennyson. 2013. Drug withdrawals and the utilization of therapeutic substitutes: The case of Vioxx. *Journal of Economic Behavior & Organization* 86, 148-168. [[Crossref](#)]
13. Anastasiya Zavyalova, Michael D. Pfarrer, Rhonda K. Reger, Debra L. Shapiro. 2012. Managing the Message: The Effects of Firm Actions and Industry Spillovers on Media Coverage Following Wrongdoing. *Academy of Management Journal* 55:5, 1079-1101. [[Crossref](#)]