Journal of Retailing 90 (2, 2014) 291–300

Journal of Retailing

The Accuracy of Scanned Prices

David M. Hardesty [a,*], Ronald C. Goodstein [b,1], Dhruv Grewal [c,2], Anthony D. Miyazaki [d,3], Praveen Kopalle [e,4]

[a] *University of Kentucky, Lexington, KY 40506, United States*
[b] *Georgetown University, Washington, DC 20057, United States*
[c] *Babson College, Babson Park, MA 02457, United States*
[d] *Florida International University, Miami, FL 33199, United States*
[e] *Dartmouth College, Hanover, NH 03755, United States*

Abstract

Nearly all retailers use price scanning systems to process transactions quickly, and presumably, accurately. Inaccuracies in scanned prices have important implications for retailers in that undercharges may impact retail profitability while overcharges could result in negative consumer perceptions of retailers and continued legal action. Here, we investigate the overall State of accuracy of scanned prices and develop certain empirical generalizations. To do so, we begin with a review of what is known regarding the accuracy of scanned prices and conduct a reanalysis of FTC data and compare the results with those from more recent cross-sectional data from the State of Washington. We then conduct a thorough analysis of a large longitudinal price scanner data set, with 231,760 products screened over a 15-year period. We investigate accuracy across different retailer characteristics, merchandise locations, and promotional activities. Implications of these results for retailers are discussed.

Keywords: Scanner data; Price accuracy; Empirical generalizations; Federal Trade Commission

Introduction

Price scanning systems have been part of the retailing landscape for over 40 years in the United States and other developed countries. Almost all retailers of major consumer goods rely on scanning systems to process consumer transactions and many offer self-scanning systems that reduce the need for cashiers. The widespread adoption of scanning systems is based on the belief that they provide more accurate prices and faster checkout times for consumers and increased profitability to retailers. These benefits, however, rely on the accuracy of these systems in terms of the match between the shelf-label price and the scanning-system's charge.

Suppose that scanned prices are systematically inaccurate. Such inaccuracies would directly affect operational efficiency, profitability, and price management for retailers. These errors would also compromise the validity of marketing models and findings that retail researchers have published utilizing these data. From the consumer perspective, scanning inaccuracies (overcharges) may result in financial losses and damage perceptions of transaction value, seller fairness, competency, and trust. These same consumer issues make the accuracy of scanned prices a continued concern for those interested in retail success and public policy, particularly during challenging economic times.

Despite the relevancy of the accuracy of scanned prices to many audiences, however, the testing of accuracy has faded in the past 15 years. The U.S. Federal Trade Commission (FTC) has, for budgetary or other reasons, shifted its efforts to other issues leaving it to weight and measures staff at the State, county, or city levels to determine if the prices scanned are accurate. These agencies have weight and measure field inspectors who

* Corresponding author. Tel.: +1 859 257 9419; fax: +1 8592573577.
E-mail addresses: david.hardesty@uky.edu (D.M. Hardesty), goodster@georgetown.edu (R.C. Goodstein), dgrewal@babson.edu (D. Grewal), miyazaki@fiu.edu (A.D. Miyazaki), Praveen.kopalle@dartmouth.edu (P. Kopalle).
[1] Tel.: +1 202 687 2009.
[2] Tel.: +1 781 239 3902.
[3] Tel.: +1 305 348 3330.
[4] Tel.: +1 603 646 3612.

http://dx.doi.org/10.1016/j.jretai.2014.03.006
0022-4359/

routinely conduct inspections funded by government budgets, while some cities (e.g., Seattle) fund it through an annual price scanning system license fee that retailers are charged. Thus, the monitoring and enforcement level has shifted and is contingent on the budgets of these various government entities. On the other hand numerous countries have begun more fervent investigations into accuracy of prices at the point of scale scanner and regulation (e.g., CCGD 2008; Government of South Australia 2008).

Although the widespread publicity surrounding price scanning accuracy in the U.S. has reduced since the 1980s and 1990s, retailers are fined for such errors to the tune of millions of dollars. For instance, in one settlement Albertsons supermarkets paid $1.85 million to resolve accusations that its scanners systematically overcharged California consumers (Briscoe 2003). In addition, and particularly concerning for academics, are recent charges that Safeway, Inc., owner of Dominick's food stores, overcharged consumers on 30% of its inspections. Dominick's is a major contributor to academic research in retailing and findings may be biased if the data were compromised (e.g., Srinivasan et al. 2004; Tsiros and Hardesty 2010).

As a result of continuing concerns regarding accuracy of scanned prices, we question whether scanning systems have made scanned prices error-free in today's environment. If not, then we suggest that marketers and retailers must adjust their models and resulting tactics to account for these errors and that issues concerning consumer trust and welfare remain relevant. We also propose that depending on the nature of the errors, if they exist, that systematic corrections might prove useful in accounting for scanner inaccuracies. Finally, we concern ourselves with the public policy implications of these findings, such as whether the FTC's exit from this arena was premature, potentially resulting in higher price inaccuracy than would have been expected with more focused regulatory intervention.

We investigate the nature of scanner inaccuracies using a longitudinal perspective across retailer characteristics, merchandise locations, and promotional activities to develop empirical generalizations regarding accuracy of scanned prices. Unlike other studies on accuracy of scanned prices, we use a large and unique data set which assessed the accuracy of 231,760 products that were checked over a 15-year period. The data were collected through 3627 store visits. This expansive data set allows us to examine accuracy across a variety of stores ranging from grocery, drug, convenience, mass merchandise, department, hardware, consumer electronics, single-brand clothing, furniture, sporting goods, books, office supply, auto parts, and other specialty stores.

We integrate insights from this large data set and a reanalysis of two secondary data sets (one from the FTC and a cross-sectional data set from Washington State) to present a number of empirical generalizations about drivers of accuracy of scanned prices. Next, we examine the FTC and Washington State data and suggest several research propositions. Then, we develop empirical generalizations based on the longitudinal data. Finally, we present a discussion of the implications of our findings.

FTC test of accuracy of scanned prices

Accuracy research has been mostly dormant for over 15 years. Like this paper, Pickering and Gaur (2009) called for retesting in their own country and found accuracy issues were still a problem in need of public policy attention in New Zealand. While the absence of research is somewhat shorter in the U.S., its implications are more widespread given the number of citizens in the U.S. and the use of U.S. data across the world of research. The last major data collection was in 1998, though these data have remained mostly unexamined (see Bergen et al. 2005 for a notable exception).

In 1995, the National Conference on Weights and Measures (NCWM) adopted the Examination Procedure for Price Verification to assess pricing accuracy in retail stores. This procedure defined inaccuracies as events in which the scanned price is unequal to the lowest price advertised, quoted, posted, or marked. The NCWM process dictated sampling procedures and called for a 98% accuracy rate for minimum compliance. Utilizing this procedure, the FTC (1998) conducted a broad-based scanner inspection that employed officials from the Department of Weights and Measures in 36 States and the U.S. Virgin Islands across a three-month period. Inspectors collected data in 1033 stores with checkout scanners, for a total of 1776 inspections (743 of the stores were visited on two occasions, once to evaluate sale items and again to focus on non-sale items). Specifically, inspectors visited and evaluated goods in 303 grocery stores, 151 drug stores, 138 department stores, 94 hardware stores, and 205 mass merchandise stores, as well as several other store types not evaluated here. In total, 107,096 goods were evaluated for accuracy of scanned prices where representatives from each State selected the stores to inspect. After the stores were selected they were subdivided via a stratified sampling procedure based on the physical retail space, and a random choice identified items from each stratum for assessment.

Testing the FTC data

The overall results from this FTC study revealed pricing errors for one of every 30 items purchased. Although the errors appeared approximately equal for overcharges and undercharges, different patterns emerged for sale items, for which errors were more frequent with an overcharge rate of approximately 67% (see also Bergen et al. 2005). Despite progress in reducing scanning errors, more errors existed than in prescanner days. Moreover, the errors were greater than previously recorded. For example, Goodstein (1994) reported that an average grocery bill of $77 contains 30 items, purchased at an average of $2.57 per item; for mass merchandise stores, the average bill was $134, with 30 items purchased at $4.47 per item. The average overcharges equaled $.24, or 6.9%, whereas average undercharges were $.83 or 23.7% (Goodstein 1994). Our study found similar sized undercharges (22% for both sale and non-sale items), but the average overcharges were 19% for non-sale items and 38% for sale items. Thus, it appears that overcharges have become more severe despite the percentage of mistakes declining.

Table 1
FTC results by store format: rate and relative size of undercharges versus overcharges.

Store type	Undercharges		Overcharges		Net effect
	Rate (%)	Relative size (%)	Rate (%)	Relative size (%)	
Grocery	1.06	19.06	1.36	23.63	.12%
versus Drug	0.91	20.57	2.11[a]	37.37[a]	.60%
versus Hardware	2.74[a]	19.28	3.49[a]	24.43	.32%
versus Department	2.58[a]	23.98[a]	1.81[a]	32.54[b]	−.03%
versus Mass Merchandise	1.79[a]	24.75[a]	1.23	34.33[a]	−.03%

[a] $p < .01$.
[b] $p < .05$.

Below we present more specific tests of the data based on the issues classified as important in earlier research studies. Based on the results of these more directed tests, we develop a series of findings supported by the FTC data.

FTC findings

Benchmarking against historical findings, Goodstein (1994) predicted that the rate of undercharges for non-sale items in stores using scanner systems should be less than the historical .7% rate that existed prior to scanner technology (Welch and Massey 1988). Instead, the undercharge rate on regular-priced, non-advertised items was 4.77% ($z = 3.91$, $p < .01$). We suggest that this rate should have fallen since Goodstein's data collection due to advanced scanner technology. The current data supported this notion; the undercharge rate for non-sale items equaled 1.94% ($z = 4.16$, $p < .01$).

FTC Finding #1: The rate of undercharges for regular-priced items has dropped since Goodstein's (1994) study.

The FTC data revealed undercharges occurred 1.94% of the time, whereas overcharges emerged only 1.22% of the time ($z = 9.71$, $p < .01$) for non-sale items. We expect that these errors are due to a lack of updating prices from their prior sale level.

FTC Finding #2: The rate of undercharges for non-sale items is greater than the rate of overcharges for non-sale items in stores that employ scanner systems.

We also evaluated each error as a percentage of the price charged for the product. The results revealed slight differences in the relative size of the errors on non-sale items. Across all store types, the undercharges were greater (21.75%) than the overcharges (19.24%; $t_{1789} = 1.75$, $p = .08$), an effect driven by greater undercharges (26.87%) compared with overcharges (19.28%) for department stores ($t_{275} = 2.93$, $p < .01$). The results suggested that undercharges are more likely and are larger for non-sale items.

FTC Finding #3: The relative size of overcharges for non-sale items is less than the relative size of undercharges for non-sale items in stores that employ scanner systems.

Pricing accuracy for sale items

The FTC data revealed that overcharges occur almost twice as often (2.28%) as undercharges (1.28%; $z = 11.23$, $p < .01$). Furthermore, when we controlled for the price of the product, the relative errors were greater for overcharges (37.62%) than for undercharges (22.07%; $z = 7.53$, $p < .01$) and these results were similar across store types. It is likely that these errors are due to poor training, inattention, or lack of staff at the individual store level.

FTC Finding #4: The rate and relative size of overcharges is greater than the relative size of undercharges for sale items in stores that employ scanner systems.

Pricing accuracy and store format

The sample used to test the remainder of the specific inquiries consisted of 96,227 items. Items from stores other than the five formats evaluated here were not considered based on the applicability, prevalence, and history of scanner checking. Store formats that were not studied prior to the FTC's data collection were excluded.

Using the Scheffé procedure to test for multiple comparisons, the differences in the percentage of errors across store formats revealed that grocery stores experienced fewer undercharges (1.06%) than did hardware stores (2.74%; $z = -12.28$, $p < .01$), department stores (2.58%; $z = -12.47$, $p < .01$), and mass merchandise stores (1.79%; $z = -7.40$, $p < .01$). No significant difference appeared between grocery stores and drugstores (.91%; $z = 1.50$, $p > .10$). The results are summarized in Table 1.

Analysis of variance revealed differences in the relative size of undercharges across store formats ($F_{4,\ 1565} = 10.25$, $p < .01$). The relative size of undercharges for grocery stores (Mean = 19.06%) was significantly smaller than for department stores (Mean = 23.98%, $p < .01$) and mass merchandise stores (Mean = 24.75%, $p < .01$) but not for drugstores (Mean = 20.57%, $p > .10$) or hardware stores (Mean = 19.28%, $p > .10$). The overall results further suggested that grocery stores and drugstores have smaller, as well as less frequent, undercharges. These results were expected, as grocery stores have received the greatest scrutiny from policy makers and should be most accurate.

Grocery stores (1.36%) were less likely than drugstores (2.11%; $z = -6.02$, $p < .01$), hardware stores (3.49%; $z = -13.78$, $p < .01$), and department stores (1.81%; $z = -3.73$, $p < .01$) to overcharge, whereas mass merchandise stores (1.23%; $z = 1.35$, $p > .10$) did not reveal significantly different rates (see Table 1). It is likely that grocery stores and mass merchandisers carry a large number of products and frequently offer sales, and as a consequence have gotten better in handling price changes. It is also likely that grocery and mass merchandise stores have better developed retail analytic capabilities and have spent more time and money in ensuring that their prices are accurate.

FTC Finding #5: The rate and relative size of undercharges is lower in grocery stores and drug stores than in hardware stores, department stores, or mass merchandise stores.

FTC Finding #6: The rate and relative size of overcharges is lower in grocery and mass merchandise stores than in drugstores, hardware stores, or department stores.

Pricing accuracy and product price

As shopping involvement increases, there is a higher likelihood that consumers scrutinize prices more carefully. Results supported this proposition (β = –.003; $\text{Wald}_1 = 13.12$, $p < .01$). As product price increases, the number of overcharges relative to undercharges declines.

FTC Finding #7: The ratio of overcharges to undercharges declines as the price of the product being purchased increases.

Discussion

The FTC data allowed us to analyze general trends in error findings since the Goodstein (1994) study. The FTC data were richer than those used in earlier studies of scanner price accuracy because they are more comprehensive and identify specific seller types in the U.S. While these results appear to be about equal across error types, recall that 20 years ago Goodstein (1994) cautioned that equal errors provide a "double whammy" for retailers. Legal actions required by law involve only overcharges and this yields large fines, penalties, and a loss of consumer patronage and trust.

Retailers' claim undercharges should be subtracted from overcharges to balance the ledger before such legal matters are decided. They argue that overcharges mean extra money taken from the consumer and undercharges represent extra money they are given back. Despite this viewpoint, one industry executive speaking with the promise of anonymity indicated that undercharges may simply be a creative fiction as many undercharges simply indicate that store staff were slow to remove expired shelf tags and that there is no economic value given to the customer in terms of real dollars.

Washington State test of accuracy of scanned prices

As mentioned earlier, the U.S. Federal Government has abandoned broad-based tests of accuracy of scanned prices. Instead, the government has asked individual States to monitor accuracy of scanned prices. One State, Washington, allowed us access to their data in order to offer a more in-depth analysis of accuracy of scanned prices. These data were used to reexamine several of the FTC findings and were based on accuracy of scanned price assessments conducted during the last quarter of 2005 across 559 stores. For each store, the data collection assessed 100 items, comparing the lowest price posted and the scanned price. Undercharges, overcharges, and the relative size of each error were identified.

Pricing accuracy for regular and sale priced items

Recall that the FTC data indicated that the relative size of overcharges for non-sale items will be less than the relative size of undercharges for non-sale items in stores that employ scanner systems. In order to test this proposition we evaluated each error as a percentage of the price charged for the product. In contrast with our findings from the FTC data, across all store types, undercharges (23.49%) and overcharges (27.03%) did not differ significantly in size ($t_{941} = -1.53$, $p > .10$). These results suggest that scanner errors for non-sale items may not be biased in the direction of larger undercharges, as suggested by the FTC data. These results also may mean good news for marketing managers; sellers may not be losing money by charging lower prices than those stated on the shelf.

Consistent with the results based on the FTC data, we found that when we controlled for the price of the product, the relative errors were greater for sale overcharges (36.39%) than for sale undercharges (26.46%; $t_{205} = 2.27$, $p < .05$). These results provided further evidence that when items are on sale, overcharges are greater than undercharges, which raises concerns for consumers in particular and may suggest the need for action by policymakers. Consumers are paying more than they should when items are promoted as being on sale. As consumers perceive such overcharges, retailers may suffer negative reactions due to scanner inaccuracies.

Consistent with the FTC data, the Washington State data also showed that as the price of the product rose, the number of overcharges relative to undercharges declined significantly (β = –.003; $\text{Wald}_1 = 4.89$, $p < .05$). That is, buyers appeared more likely to identify overcharges when the price of the good was high rather than low. These results suggest that policymakers initially should focus on lower-priced goods in their monitoring efforts.

The Washington State data were useful in providing greater confidence in several of the FTC data findings. In the next section, we discuss a more comprehensive analysis of accuracy of scanned prices using a longitudinal dataset covering 15 years and including information related to retailer characteristics, merchandise locations, and promotional activities.

A longitudinal assessment of accuracy of scanned prices

In order to obtain a database that had both currency and breadth, retail accuracy of scanned prices data collected from 1996 to 2010 was obtained from the State of Washington's Consumer Affairs Unit. Longitudinal analysis has the benefit of being able to identify trends across time. This is a large data set comprised of 3628 screenings conducted over 15 years (about 231,760 observations). Each screening typically screened between 50 and 100 products to determine the overall number of errors, overcharges, and undercharges. We divided the number of errors by the number of items scanned to create percentages associated with overall errors, overcharges, and undercharges (see Appendix A). The data did not include the level of overcharges or undercharges. In addition to accuracy of scanned prices assessments, the data account for a variety of retail factors including

Table 2
Descriptive statistics for longitudinal data.

Independent variable	Mean	Standard deviation
Percent from regular shelves	.86	.32
Percent seasonal items	.02	.04
Percent end-of-aisle items	.07	.06
Percent of items at check out	.03	.05
Percent of items checked with PLU identifiers	.20	.38
Percent direct store delivery items	.06	.11
Percent items on temporary price reduction	.08	.12
Percent items advertised in flyer	.02	.08
Percent of items with permanent discounts	.04	.09
Percent with store coupons	.00	.01

information regarding retailer characteristics, merchandise location, and promotional activity.

With respect to retailer characteristics, we coded the data into retailer types which have had greater regulation of accuracy of scanned prices versus those that have had less. This retailer regulation factor was classified based on the authors' reviews of the prior studies presented in the marketing and retail literature (e.g., Clodfelter 1998). We specifically focus on four retailer types which have been heavily regulated for accuracy of scanned prices (grocery: $n = 740$, drug: $n = 498$, convenience: $n = 376$, and mass merchandise stores: $n = 87$) and combine all other retailer types which have not been as heavily regulated (department: $n = 117$, hardware: $n = 210$, consumer electronics: $n = 186$, single-brand clothing: $n = 330$, warehouse: $n = 17$, sporting goods: $n = 120$, books: $n = 50$, office supply: $n = 151$, auto parts: $n = 122$, other specialty: $n = 155$). The data also include dichotomous information regarding the number of stores on a particular chain (Small = 3 or fewer locations, Large = 4 or more locations).

In addition, the data include several factors related to merchandise location. Specifically, information regarding the number of items checked from regular shelves, check-out stands, end-cap displays, SKU or PLU (price look up) code identifiers (often used in produce departments), as well as the number of seasonal items (often located in a specific part of a retail outlet) and items delivered directly to the store. Each of these was converted to percentages to account for the number of items scanned for each screening.

Finally, the data provide information regarding the promotional activity associated with each scanned product. Information regarding the number of items on temporary price reduction, advertised, permanently discounted, and in store coupons was included. Table 2 includes descriptive statistics for each independent variable.

Below, we provide two sets of analyses. The first are descriptive analyses of the scanner error percentages. Then, we provide regression analyses for each of our dependent variables including overall error percent, overcharges and undercharges.

Descriptive analyses

As displayed in Table 3, the overall error percentage was 4.08% suggesting a significant problem associated with scanner inaccuracy as FTC standards for inspections suggest fewer than two percent errors. The average is much higher than the two percent regulation put forth by the FTC ($t_{3626} = 21.46$, $p < .01$) as well as the industry standard of .70% ($t_{3626} = 34.85$, $p < .01$). Of these errors, 1.75% of them are overcharges that could result in both lawsuits and negative attitudes toward retailers. Additionally, 2.33% of the errors are undercharges and either negatively impact retailer profitability if real or reflect their lack of employee timeliness in updating shelf tags. In percentage terms, there are significantly more scanner errors that are undercharges ($t_{3626} = 8.10$, $p < .01$). These results lead to the following empirical generalizations:

Table 3
Longitudinal data average error percentages.

Error type	Average	Standard deviation
Overall error percentage	4.08	5.85
Overcharges	1.75	3.43
Undercharges	2.33	3.83

EG1: Average scanner error percentages significantly exceed the recommended level of two percent suggested by the FTC and the industry standard of .70%.

EG2: There are significantly more undercharges than overcharges in stores employing UPC scanners.

Regression models

We next developed three regression models to predict the overall error percentage and level of overcharges and undercharges, respectively. For each regression, we also included the independent variables pertaining to retailer characteristics, merchandise location, and promotional activity (see Appendix A). Finally, each regression included an independent variable capturing the year the screening took place (Year).

Table 4 presents the results associated with the regression models. The models are significant for the overall error percentage ($F_{13,\ 3603} = 27.73$, $p < .01$), overcharges ($F_{13,\ 3603} = 14.39$, $p < .01$), and undercharges ($F_{13,3603} = 27.81$, $p < .01$), explaining nine, five, and nine-percent of the variance inaccuracy of scanned prices, respectively. In order for an empirical generalization to be warranted, the overall error percentage and at least one of the other error percentages (overcharges or undercharges) needed to be significant. When these criteria were not met, we reported the results without establishing an empirical generalization.

Retailer characteristics

Past research has examined a host of retail or store specific variables, such as retail square footage, sale footage, and shelf space (e.g., Gauri 2013; Grewal et al. 1999). We examine two retail characteristics that were available in our data set, namely the number of stores in the chain and whether the retailer operates in an industry which has been regulated for scanner accuracies.

Our dichotomous measure of the number of stores in the retail chain impacted errors, while the extent to which the retail industry has been regulated did not. Retailers with more outlets had

Table 4
Error percentage regression results.

Variable	Overall errors		Overcharges		Undercharges		VIFs
	Estimated coefficient	Standard error	Estimated coefficient	Standard error	Estimated coefficient	Standard error	
Intercept	3.446**	.525	.316	.314	3.138**	.344	
Retail							
Retailer type	−.003	.003	.001	.002	−.004*	.002	2.00
Number of stores in chain	−.018**	.002	−.009**	.001	−.010**	.001	1.50
Merchandise							
Percent regular shelves	.003	.003	.001	.002	.001	.002	1.03
Percent seasonal items	−.051*	.026	−.021	.015	−.031+	.017	1.12
Percent end-of-aisle items	.011	.016	.011	.010	.000	.010	1.07
Percent of items at check out	.069**	.019	−.008	.011	.077**	.012	1.14
Percent of items checked with PLU identifiers	.017**	.003	.003+	.002	.013**	.002	1.31
Percent direct store delivery items	.093**	.011	.034**	.007	.060**	.007	1.66
Promotion							
Percent items on temporary price reduction	−.033**	.009	−.028**	.005	−.004	.006	1.19
Percent items advertised in flyer	.039**	.012	.022**	.007	.017*	.008	1.13
Percent of items with permanent discounts	.012	.011	−.017**	.006	.030**	.007	1.15
Percent with store coupons	−.149	.109	−.078	.065	−.072	.071	1.03
Year	−.002**	.000	.000	.000	−.002**	.000	1.26
R-square	0.09		0.05		0.09		
Number of Observations	3617		3617		3617		

** Significant at .01 level.
* Significant at .05 level.
+ Significant at .10 level.

significantly fewer errors ($\beta = -.018$, $p < .01$), indicating that retailers with fewer outlets may need to monitor accuracy of scanned prices even more closely. This effect held true for both overcharges ($\beta = -.009$, $p < .01$) and undercharges ($\beta = -.010$, $p < .01$). It is likely that retailers with more outlets have fewer errors as they have devoted more attention to the management of price scanning systems at the store level. Some of these retailers have internal audit teams that help ensure their prices are accurate. Thus,

EG3: Average scanner error percentages are significantly higher for retailers with fewer outlets versus retailers with more outlets.

Retail stores which historically have been more heavily regulated have the same overall error percentages ($\beta = -.003$, $p > .10$) and overcharges ($\beta = .001$, $p > .10$). However, undercharge rates are lower for more regulated industries compared to those who have been less regulated ($\beta = -.004$, $p < .05$) suggesting that less regulated stores need to improve their procedures for loss or shelf updating.

Merchandise characteristics

Past research has examined the role of various merchandise characteristics sparingly, with the exception of whether the item was on display (e.g., Goodstein 1994; NIST 1993). We examined the role of a number of merchandise variables including whether the items checked came from regular shelves, check-out stands, end-cap displays, items requiring SKU or PLU entry, seasonal items (often located in a specific part of a retail outlet), and items delivered directly to the store.

The percentage of seasonal items, check stand displays, items checked via SKU or PLU, and the percentage of items delivered directly to the store each significantly impacted error percentages. The percentage of regular shelf items and end cap displays did not significantly impact error percentages. As the percentage of seasonal items rises, the overall error percentage goes down significantly ($\beta = -.051$, $p < .05$) as does undercharges ($\beta = -.051$, $p < .05$), indicating that seasonal items appear well-tracked by retailers. Conversely, the overall error percentage rises significantly as the percentage of items on check-out display increases ($\beta = .069$, $p < .01$), PLU code identifiers ($\beta = .017$, $p < .01$), and direct store delivery ($\beta = .093$, $p < .01$). For check-out displays, undercharges increase significantly ($\beta = .077$, $p < .01$). For PLU code identifiers, both overcharges ($\beta = .003$, $p < .10$) and undercharges ($\beta = .013$, $p < .01$) increase significantly. The same pattern is found for direct store delivery in that both overcharges ($\beta = .034$, $p < .01$) and undercharges ($\beta = .060$, $p < .01$) rise. These results suggest that retailers should be aware of scanner inaccuracies associated with items offered at check-out displays, items coming from areas of the stores where perishable goods are purchased (e.g., produce departments) and PLU code identifiers are used, and for items delivered directly to the store by the manufacturer. Thus,

EG4: When retailers offer seasonal versus non-seasonal items, average scanner error percentages are significantly lower.

EG5: When retailers (a) offer items on check-out displays, (b) use PLU identifiers or (c) pertain to direct store delivery products versus when they do not, there are significantly more scanner errors.

Table 5
Error percentage regression results including demographics.

Variable	Overall errors		Overcharges		Undercharges		VIFs
	Estimated coefficient	Std. error	Estimated coefficient	Std. error	Estimated coefficient	Std. error	
Intercept	3.282**	.523	.215	.312	3.077**	.344	
Retail							
Retailer type	−.003	.003	.001	.002	−.004*	.002	2.08
Number of stores in chain	−.018**	.002	−.009**	.001	−.009**	.001	1.51
Merchandise							
Percent regular shelves	.002	.003	.001	.002	.001	.002	1.03
Percent seasonal items	−.053*	.026	−.020	.015	−.033*	.017	1.13
Percent end-of-aisle items	.008	.016	.008	.010	−.001	.010	1.08
Percent of items at check out	.072**	.019	−.006	.011	.077**	.012	1.14
Percent of items checked with PLU identifiers	.016**	.003	.003+	.002	.013**	.002	1.32
Percent direct store delivery items	.089**	.011	.030**	.007	.059**	.007	1.69
Promotion							
Percent items on temporary price reduction	−.030**	.009	−.027**	.005	−.003	.006	1.20
Percent items advertised in flyer	.038**	.012	.022**	.007	.017*	.008	1.13
Percent of items with permanent discounts	.018+	.011	−.013*	.006	.031**	.007	1.18
Percent with store coupons	−.134	.108	−.068	.065	−.066	.071	1.03
Year	−.002**	.000	.000	.000	−.002**	.000	1.27
Zip code demographics							
Median income	.000**	.000	.000**	.000	.000	.000	3.20
% with college degree	−.001**	.000	.000**	.000	.000**	.000	3.42
Population	.000*	.000	.000*	.060	.000+	.000	2.66
Average household size	.002	.003	.000	.002	.003	.002	2.25
Median age	.000	.000	.000	.000	.000	.000	4.09
% Married	−.001**	.000	.000**	.000	.000+	.000	3.69
R-square	0.10		0.06		0.09		
Number of observations	3617		3617		3617		

** Significant at .01 level.
* Significant at .05 level.
+ Significant at .10 level.

There are no significant effects on overall error percentages for the percentage of items checked from regular shelves ($\beta = .003$, $p > .10$) or the percentage of items from end-cap displays ($\beta = .011$, $p > .10$). Similarly, there were no significant differences ($p > .10$) in overcharge and undercharge rates within these categories.

Promotional characteristics

Past behavioral research consistently demonstrates the importance of price promotions on perceptions of value and behaviors (e.g., Biswas et al. 2013; Puccinelli et al. 2013). Similarly, panel based studies systematically have examined the role of these promotional variables on shopping behaviors (e.g., Ma et al. 2011).

Both items on temporary price reduction and advertised sale items impacted accuracy of scanned prices, while permanently discounted items and items affiliated with store coupons did not. As the percentage of items on temporary price reduction increased, the overall error percentages ($\beta = -.033$, $p < .01$) and overcharges significantly declined ($\beta = -.028$, $p < .01$). However, advertised sale items were associated with significantly greater error overall ($\beta = .039$, $p < .01$), overcharge percentages ($\beta = .022$, $p < .01$), and undercharge percentages ($\beta = .017$, $p < .05$). These result in the following empirical generalizations:

EG6: When items are on temporary price reduction, errors are significantly less likely.

EG7: When sale items are advertised, errors are significantly more likely.

Items on permanent discount did not impact overall error percentages ($\beta = .012$, $p > .10$). Similarly, items affiliated with store coupons did not significantly impact overall error percentages ($\beta = -.149$, $p > .10$), overcharges ($\beta = -.078$, $p > .10$), or undercharges ($\beta = -.072$, $p > .10$).

Supplemental analysis

A number of retailing studies incorporate the effects of customer demographics into their analysis (e.g., Gauri, Trivedi and Grewal 2008; Talukdar, Gauri and Grewal 2010). Thus, in addition to the analyses above, an analysis of demographic information of the individuals in each retailer zip code was conducted. These analyses included an exploration of the effects of income, education, population, household size, age, and percentage married on overall scanner errors, overcharges, and undercharges. The results of these analyses are included in Table 5. The results reveal that each of the seven empirical generalizations remains stable when accounting for these individual demographics. Moreover, income and population are associated with greater overall scanner errors. These results are consistent with those of Goodstein (1994) in that retailers make more errors in more populated and affluent locations. In addition, when the individuals in a location have a higher percentage of college degrees and a higher percentage married, overall scanner errors are lower.

The same pattern of effects emerged for overcharges in that they were more likely for high income locations and more populated areas while they were less likely for more highly educated areas and those areas with more married individuals. Undercharges were not significantly more likely for locations with higher income individuals ($p = .14$) and only marginally higher for areas with greater population ($p = .08$). Finally, undercharges are more prevalent for more educated areas and for locations with more married individuals.

Discussion and implications

In the current marketplace, scanner systems are employed across almost every retail format. Compared to their adoption rate, the testing of accuracy of scanned prices has not kept pace. If these systems are not as accurate as promised then implications related to their operational efficiency, profitability, and price management for retailers must be called into question. The same caution must be accorded toward academic research that has utilized the data ascertained from these systems, especially where that research incorporates promotional advertising effects. In this research, we conducted several analyses of accuracy of scanned prices including a broad-based test utilizing FTC data, retesting of several FTC findings using more current data from Washington State, and finally the development of empirical generalizations based on analysis of a large longitudinal study made available by the Washington State Office of Consumer Affairs.

There are individual limitations associated with each of these data sets that should be considered before drawing final conclusions on these matters. For instance, the FTC data were collected some time ago and the longitudinal data did not include the specific size of each retail location or the level of overcharges or undercharges. Despite these limitations, our findings hold interesting and important implications for retail practitioners, public policymakers, and academic researchers.

Implications for retailers

Overall, accuracy of scanned prices remains an issue within retail stores. Though error rates have improved, retailers would be well-advised to scrutinize the accuracy of their price scanning activities more carefully because the rate of errors still far exceeds the FTC standard of 2% and industry standard of .70%. Undercharge rates continue to plague the systems, which may just be poor management of shelf-tag systems at best, or lost profitability at worst. In either case, retailer attempts to estimate price elasticities will be in error due to inaccuracies between consumer and retailer perceptions of the price being charged.

Scanner errors remain a vital concern to the consumers that patronize retail stores, just as they were decades ago. We conducted an additional survey to better understand how consumers feel with respect to scanners and scanner errors. The survey made use of a convenience sample recruited from two grocery store chains located in a major U.S. city ($n = 174$: 53% women and average age 51.5 years). The results indicated that shoppers strongly agreed that relative to item pricing, scanners significantly reduce time (Mean = 6.29/7, $p < .01$) and provide more information (Mean = 5.74/7, $p < .01$), though the absence of item prices makes accuracy of scanned prices significantly harder to check (Mean = 5.12/7, $p < .01$). Buyers expressed significantly ($p < .05$) decreased attitudes, lower patronage intentions, and more negative word-of-mouth intentions when they were aware that retailers overcharged items in the store. These findings illustrate that consumer discovery of overcharging has relevant implications for retailers with respect to brand attitude and sales.

Our merchandise characteristics tests indicate lower scanner error percentages for retailers with more stores in the chain versus those with fewer outlets, suggesting that retailers with fewer outlets need to be even more concerned with scanner errors. Further, when retailers offer seasonal versus non-seasonal items, average scanner error percentages are less problematic suggesting that the diligence associated with the accuracy of prices affiliated with seasonal items be accorded to the rest of the inventory as well. Similar diligence is due for items offered on check-out displays, those utilizing PLU identifiers (e.g., produce), and those delivered directly to the store (e.g., breads and baked items), which often are stocked by outside vendors.

Analyses of promotional activities indicated that temporary in-store sales promotions are not a significant source of additional errors, but advertised sales specials are. Interestingly, it is this latter category of errors that has been of the most concern in scanner studies throughout the past two decades (e.g., NIST 1993). Finally, although errors have become less problematic across time, error percentages remain relatively large compared to FTC standards.

Public policy implications

Each result that pertains to the retailer has a corollary for the public policy side of the ledger, as policymakers are urged to investigate accuracy rates and directionality of scanned price violations when errors are ubiquitous. Increased regulation and attention to grocery stores appears to have improved accuracy of scanned prices based on the FTC analyses as scanner errors have decreased over time. Similar improvements were found in other global regions employing these systems (e.g., Pickering and Gaur 2009). Yet domestic and foreign investigators caution that ignoring these issues will result in regulatory violations and potentially increased error rates. Some specific retail formats, such as hardware stores, may need greater regulation in that overcharges are more prevalent and relatively large in these formats. Regulatory focus and testing appear to increase retailers' vigilance in assuring accuracy of scanned prices and are likely to ensure that retailers take more proactive roles to avoid lawsuits and negative publicity (Briscoe 2003).

As a potential way to protect against negative consumer outcomes associated with scanner inaccuracy, The Retail Council of Canada created a voluntary program called the Scanner Price Accuracy Voluntary Code (http://www.competitionbureau.gc.ca/eic/site/cb-bc.nsf/eng/01262.html; 2002). This code protects consumers for up to $10 per scanner overcharge identified. A similar code could be developed in the U.S. to protect consumers against overcharges, even though many retailers offer their own scanner accuracy guarantee policies. We strongly advise consumers to ask about such policies when errors are discovered.

It appears that consumer complaints of the inaccuracy about prices scanned for package goods are generally valid as they have evidence such as receipts and advertised prices. Now armed with smart phones and apps that make them aware of price promotions (Grewal et al. 2012), consumers have the ability to photograph the price at the shelf or display. Thus, complaints may be taken more seriously and enforced in a fashion wherein the violating retailer is cited, fined, and compelled to ensure to compensate for any overcharges made.

Particularly concerning is the fact that we found significant difference in error rates based on zip codes. Although unreported in his earlier work, Goodstein (1994) found significantly more errors in wealthy neighborhoods compared to error rates (within the same chains) in poorer neighborhoods. This result may imply that retailers might be purposely managing errors such that stores had fewer overcharges in areas where residents were more sensitive to prices and more in areas where consumers might be less price-sensitive. Our zip code analysis echoes this concern in that errors were higher in more affluent neighborhoods relative to error rates in less affluent neighborhoods. While we cannot speak to causality, together the results might be worth public policy investigation.

A lot of the funding for weights and measures programs comes from fees for devices like scales and gas pumps and as a consequence a disproportionate amount of attention is devoted to these types of products. Weights and measure programs should increase the emphasis and oversight devoted to packaged goods since there are numerous concerns in addition to inaccuracies of scanned price, such as labeling and content (net weight) issues.

Implications for researchers

Overall, there has been progress in reducing scanning errors in the retail environment, especially in supermarkets whose data most often are utilized in marketing and retail research. Real concern, however, remains that more errors exist today than in prescanner days and that the pattern of these errors depend on retailer type, promotional activities, and merchandise location decisions. If such error patterns are not considered, there is a strong probability that the magnitude of the effects reported in research is misguided. Incorporating these biases in the data may yield more accurate descriptive conclusions and normative recommendations. This suggestion is similar in spirit to the joint scanner panel/choice experiment model suggested by Swait and Andrews (2003) where scanner panel data are supplemented with choice experiment data to improve predictions. Unit sales figures may provide more accurate representations of the effects of marketing tactics compared to dollar sales when scanner inaccuracies are suspected.

In conclusion, the issue of scanner accuracy has been relatively dormant for well over a decade. Errors still exist at a significant level, however, and improvement of errors rates is not the same as the elimination of such errors. Further, new scanner forms are entering into the market that are purported to provide more ease and accuracy in pricing, such as EAN Codes, GS1, and RF tags. While these formats are based on the original bar scan format, they are more detailed and their impact should be tested as well. The accuracy of scanners affects many audiences including retailers, consumers, policymakers, and researchers. The broad and continuing expansion of scanning systems to new markets dictates that attention should be given to testing these systems across multiple retail settings.

Appendix A. Key measures for longitudinal study

Dependent variables

- ERROR PERCENTAGE – the total number of errors/number of items scanned
- OVERCHARGES – the number of overcharge errors/number of items scanned
- UNDERCHARGES – the number of undercharge errors/number of items scanned

Independent variables in regression analysis

- Retail characteristics
  - ◯ RETAILER TYPE – coded as “1” if the retailer is heavily monitored (grocery, drug, convenience, and mass merchandise) and “0” if the retailer has not been heavily monitored (department, hardware, consumer electronics single-brand clothing, furniture, sporting goods, books, office supply, auto parts, other specialty)
  - ◯ NUMBER OF STORES IN CHAIN – coded as “1” if the retailer has four or more locations and “0” if the retailer has three or fewer locations
- Merchandise location
  - ◯ The percentage of items from regular shelves
  - ◯ The percentage of items which were seasonal
  - ◯ The percentage of items from end cap displays
  - ◯ The percentage of items from a check stand display
  - ◯ The percentage of items checked via SKU or PLU code identifiers
  - ◯ The percentage of items delivered directly to the store
- Promotional activity
  - ◯ The percentage of items on temporary price reduction
  - ◯ The percentage of items advertised in flyer
  - ◯ The percentage of items on permanent discount
  - ◯ The percentage of items available on store coupon
- Year of data
  - ◯ The year the screening took place

References

Bergen, Mark, Daniel Levy, Sourav Ray, Paul H. Rubin and Benjamin Zeliger (2005), “When Little Things Mean a Lot: On the Inefficiency of Item Pricing Laws,” in *American Law and Economics Association Annual Meetings Paper 34*, Biswas, Abhijit, Sandeep Bhowmick, Abhijit Guha and Dhruv Grewal (2013), "Consumer Evaluation of Sale Priced: Role of the Subtraction Principle," *Journal of Marketing,* 77 (July), 49–66.

Briscoe, Daren (2003, September 3), "Scanner Errors Cost Albertsons Stiff Fine," *Los Angeles Times,*, B1.

CCGD (2008), *Scanner Price Accuracy Voluntary Code: Associations' Annual Report to the Competition Bureau,* Canadian Council on Grocery Distributors, November. http://www.ccgd.ca/home/en/PDF/National/Scanner%20Accuracy%20Report%20-%202008.pdf (accessed 05.02.09)

Clodfelter, Guy Richard (1998), "Pricing Accuracy at Grocery Stores and Other Retail Stores Using Scanners," *International Journal of Retail and Distribution Management,* 26 (11), 412–20.

Federal Trade Commission, FTC (1998), *Price Check II: A Follow-up Report on the Accuracy of Checkout Scanner Prices,* http://www.ftc.gov/reports/scanner2/scanner2.shtm

Gauri, D.K. (2013), "Benchmarking Retail Productivity Considering Retail Pricing and Format Strategy," *Journal of Retailing,* 89 (1), 1–14.

Gauri, Dinesh K., Minakshi Trivedi and Dhruv Grewal (2008), "Understanding the Determinants of Retail Strategy: An Empirical Analysis," *Journal of Retailing,* 84 (3), 256–67.

Goodstein, Ronald C. (1994), "UPC Scanner Pricing Systems: Are They Accurate?," *Journal of Marketing,* 58 (April), 20–30.

Government of South Australia (2008), *Stores Checked for Price Scanning Accuracy,* Office of Premier of South Australia, April 11. http://www.ministers.sa.gov.au/news.php?id=2996 (accessed 05.02.09)

Grewal, Dhruv, Michael Levy, Anuj Mehrotra and Arun Sharma (1999), "Planning Merchandising Decisions to Account for Regional and Product Assortment Differences," *Journal of Retailing,* 75 (3), 405–24.

Grewal, Dhruv, Anne L. Roggeveen, Larry D. Compeau, Michael Levy and Compeau Michael Levy (2012), "Retail Value-Based Pricing Strategies: New Times, New Technologies, New Consumers," *Journal of Retailing,* 88 (1), 1–6.

Ma, Yu, Kusum L. Ailawadi, Dinesh Gauri and Dhruv Grewal (2011), "An Empirical Investigation of the Impact of Gasoline Prices on Grocery Shopping Behavior," *Journal of Marketing,* 75 (March), 18–35.

National Institute of Standards and Technology (1993), *Examination Procedure for Price Verification in Food Stores,* Gaithersburg, MD: National Institute of Standards and Technology.

Pickering, Paul and Sanjaya S. Gaur (2009), "The Incidence of Scanner Fraud at the Grocery Checkout: Does Anyone Care?," *Asia Pacific Advances in Consumer Research,* 8, 357–8.

Puccinelli, Nancy, Rajesh Chandrashekeran, Dhruv Grewal and Rajneesh Suri (2013), "Are Men Seduced by Red? The Effect of Red vs. Black Prices on Price Perceptions," *Journal of Retailing,* 89 (2), 115–25.

Srinivasan, Shuba, Koen Pauwels, Dominique M. Hanssens and Marnik G. Dekimpe (2004), "Do Promotions Benefit Manufacturers, Retailers, or Both?," *Management Science,* 50 (May), 617–29.

Swait, Joffre and Rick L. Andrews (2003), "Enriching Scanner Panel Models with Choice Experiments," *Marketing Science,* 22 (4), 442–60.

Talukdar, Debabrata, Dinesh K. Gauri and Dhruv Grewal (2010), "An Empirical Analysis of Extreme Cherry Picking Behavior of Consumers in the Frequently Purchased Goods Market," *Journal of Retailing,* 86 (4), 336–54.

Tsiros, Michael and David M. Hardesty (2010), "Ending a Price Promotion: Retracting It in One Step or Phasing It Out Gradually," *Journal of Marketing,* 74 (1), 49–64.

Welch, Joe L. and T.K. Massey Jr. (1988), "Consumer cost implications of reducing item omission errors in retail optical scanner environments," *Akron Business and Economic Review,* 19 (Summer), 97–105.