Change at the Checkout:
Tracing the Impact of a Process Innovation

Emek Basker
University of Missouri

March 2013

Abstract

Barcode scanners, a foundational process innovation in the grocery supply chain, had large effects on grocery prices. I use difference-in-difference specifications with city-level price data from 1972 to 1984 and show that prices of groceries fell, on average, by about 0.75% by the time supermarkets’ adoption rate reached 5%. Several specification tests confirm that the estimates are causal. The results are consistent with prior estimates of labor saving by scanners. A conservative estimate suggests that the short-run consumer-surplus gain from scanners was approximately $3.25 billion per year in 2012 dollars.

JEL Codes: L81, D22, O33

Keywords: Barcode scanners, Retail, Supermarkets, Technology, Prices, Process Innovation

*Preliminary and incomplete. Comments very welcome to emek@missouri.edu. This paper was written while I was a visitor at Boston College and at the Stanford Institute for Economic Policy Research; I thank both institutions for their hospitality. I also thank John Meisel and Sue Wilkinson for help obtaining the Food Marketing Institute reports, C.Y. Choi for help with ACCRA data, and David Autor, Jonathan Beck, Tim Bresnahan, Cory Koedel, Saul Lach, Mark Lewis, Jeff Milyo, Petra Moser, and seminar participants at the Universities of Arizona and Oregon for helpful comments and conversations. All remaining errors and omissions are my own.
1 Introduction

Process innovations are a key to technological change at both a micro and a macro level. Yet, unlike product innovations, in the form of both new products and quality enhancements to existing products, process innovations are still poorly understood. These innovations are often invisible, as they may have no observable effect on the final product delivered to consumers, but they represent transformations of firm organization, the production process, and the supply chain. Process innovations have real economic effects by reducing the cost of production, causing resource reallocation and industrial churn, changing demand for inputs, and reducing the prices that consumers pay for final goods.

This paper provides evidence of the price effect of barcode scanning, a foundational process innovation in the retail sector. Scanning transformed the grocery sector by allowing supermarkets to encode prices in machine-readable format. This seemingly minor innovation formed the foundation for many later improvements in supply-chain management, from electronic data interchange (EDI) and electronic payments to inventory management and self-scanning. The first scanner came on-line in June 1974 at a Marsh supermarket in Troy, Ohio. By January 1985, 29% of supermarkets in the U.S. were using the technology.¹

Several features of scanning make it a good candidate for studying the impact of process innovation on prices. First, the Universal Product Code was designed as an industry standard to preempt individual supermarket chains’ introducing their own proprietary codes. It was hoped that encoding prices in a machine-readable code would speed up the checkout process and reduce its error rate, but it was not designed with consumer prices in mind, and in that sense represents an exogenous shock to prices. Second, it was a “pure” process innovation, one that had no impact on the physical product received by consumers. This contrasts with

innovations such as cotton spinning or automobile production lines, which also affected the final product. Third, the gradual diffusion of scanning across supermarkets provides natural variation in the proportion of scanning supermarkets across cities and time, which I exploit to identify the price effect of scanners. Finally, consumer prices of groceries are readily available and cleanly measured.

I find that grocery prices fell considerably in the first decade of scanning. In a difference-in-difference specification with city-level price data, a scanner penetration rate of 5% — the median in my sample for cities with at least one scanner — is associated with grocery-price decreases of about 0.75%. Back-of-the-envelope calculations suggest that the short-run gain in consumer surplus due to scanners is on the order of $3.25 billion annually in 2012 dollars. I perform several “placebo” regressions to verify that the results are not driven by endogenous scanner adoption at the supermarket level.

The most likely mechanism through which this process innovation affected prices is by reducing labor costs. Before the UPC, each can, jar, box, and bag on self-service supermarket shelves had to be individually tagged with a price sticker; cash registers were simply adding machines. Using the fact that some states required that stores continue “item pricing” even after installing barcode scanners, while other states allowed stores to switch to “shelf scanning” — which entailed placing just one sticker at the shelf, rather than the item, level — I find evidence consistent with shelf-pricing providing a significant cost saving. In addition, scanning was faster than manually hand-entering prices. An early U.S. Department of Agriculture predicted that scanners could increase checker speed by as much as 18–19%. These two effects combined to reduce labor costs in scanning stores. In earlier work (Basker, 2012), I found that scanning stores reduced their wage bills by about 4.5%.

Another potential source of saving is the early implementation of a complementary technology: price look-up codes, which reduced the need for cashiers to memorize prices, particularly for products with volatile prices. This complementarity may explain why the largest observed price effects of scanners were for produce and meat — products unlikely
to bear barcodes, but also not typically tagged with price stickers. Prior to the adoption of look-up codes, cashiers would have had to either memorize prices of produce and meat or look them up in cumbersome tables, wasting valuable time. In contrast, looking up the prices of canned goods was relatively fast even before scanning thanks to item-price stickers.

The effect I estimate is in line with the price effects of other interventions in the food market. For example, Lach (2007) estimates that a one-percentage-point increase in the ratio of recent immigrants (from the Former Soviet Union to Israel) to natives in a city decreased prices by an average 0.5%, an effect he attributes to immigrants’ lower search costs driving down the competitive price level. Basker and Noel (2009) estimate that entry by Wal-Mart Supercenters drives down incumbent supermarkets’ prices in the same city by about 1%. What is different in this paper is that the effect comes from a technological advance in the supply chain, not from external pressure from consumers or competitors.

The rest of the paper is organized as follows. Section 2 provides background on the technology and the institutional and legal environment. Section 3 describes the data sources used in this study. Section 4 includes the main results and specification tests. Section 5 concludes by evaluating the welfare implications of the finding.

2 Background and Empirical Approach

In the beginning, there was the abacus.

The mechanical cash register, patented in 1879, was the first device to print out purchase records (Brat and Zimmerman, 2009). Several innovations improved on this early idea, starting with the electronic cash register (ECR), an electronic equivalent of the manual cash register; the inter-connected ECR, which could share subtotals but had minimal data-storage ability; and the mini-computer ECR, which was able to process check authorizations and handle a large price-lookup table (Selmeier, 2008).
The barcode, or Universal Product Code (UPC), originates with Wallace Flint’s 1932 Harvard Master’s thesis. Norman Joseph Woodland and Bernard Silver applied for a patent on a similar innovation in 1949. Implementation took more than two more decades and an organization, the Ad Hoc Committee on a Uniform Grocery Product Identification Code, that brought together food manufacturers, wholesalers, and retailers, with the help of consulting company McKinsey & Co., to agree on a standard. Food manufacturers in particular were in a hurry to create an industry standard to preempt individual supermarket chains’ introducing their own proprietary codes, which would have created a burden on manufacturers. By this point, Wallace Flint was himself a Vice President of the National Association of Food Chains, and Norman Joseph Woodland was employed as an engineer by IBM and was a leader of IBM’s efforts to get into the checkout business. The Ad Hoc Committee included representatives of one independent grocer and one cooperative, along with executives of the largest supermarket chains: Kroger, A&P, and Super Value (Haberman, 2001, p. 145).

The early UPC code consisted of 12 digits: a “number system” digit, a 5-digit vendor identification code, a 5-digit item code, and a check digit. The vendor code was assigned by a central organization to member companies. Each vendor-member had the ability to assign up to 100,000 item codes to individual items. As of July 1974, the Uniform Grocery Product Code Council (UGPCC) had about 1,500 members, most of them food manufacturers; only members could print the UGPC (later, UPC) on their products, but not every member printed a barcode on every package, as redesigning package labels and upgrading printers took time. Absent detailed information on label redesign, it is impossible to know which packaged products bore barcodes at any given date. The check digit ensured that no more than one in ten errors in scanning would be allowed to go through. The number-system digit was “0” for regular grocery products and “2” for random- (or variable-)weight products.

---

2 The organization was initially the Distribution Number Bank (DNB), then Distribution Codes Incorporated (DCI) (Selmeier, 2008, p. 110).

3 The check digit was calculated as the difference between ten and the last digit of the sum of the odd-number digits multiplied by three plus the sum of the even-numbered digits.
packaged in store.⁴ For these products, such as meat, deli, and some produce, the entire code save the number system and the check digit was assigned in-store, and the last four digits designated the price, up to $99.99 (Selmeier, 2008, pp. 111, 121).

Scanning stores reduced their labor costs by approximately 4.5% (Basker, 2012). There are two main channels through which this effect could have taken place: labor saving at the cash register, due to faster checkout times and/or the ability to use lower-skilled workers (or to reduce training time for new cashiers); and labor saving for price changes on the store floor.

Several features of scanning/electronic cash registers may make them faster than manual price entry, especially after a bit of practice. The UPC code is printed in a fixed location on the product label, and saves looking at the product for the price sticker. Even for items that may not have been barcoded early on, such as variable-weight products (produce, meat), a fixed code that could be memorized saved the cashier the need to memorize new prices each time the price was changed. A store manager remarked to Progressive Grocer magazine in 1976 that “today the turnover on the front end is so fierce that training the girls [sic] to tell a Rome from a Delicious apple, or a Texas from a Florida grapefruit, is time consuming. Worse, it’s pretty ineffective. On the other hand, they quickly memorize most of the codes” (Progressive Grocer, March 1976, pp. 40-42). Consistent with this explanation, U.S. Department of Agriculture predictions, reported by Bloom (1972, pp. 218-220), indicated that if 100% of items were to be barcoded, checker speed could increase by as much as 18–19%; the productivity gain would have been lower in the early years, with only partial adoption, but could still have been substantial. Newspaper reports, as well as informal interviews I conducted with store managers about their personal experiences with scanner adoption, confirm that staffing levels were reduced in scanning stores (see, e.g., Williams, 1984).

⁴Further allowed values were “3” for drugs and “5” for coupons.
In addition, some scanning stores were able to use the new technology to eliminate the need to place a price sticker on each and every item in stock. The availability of scannable UPCs, and of time-invariant codes for variable-weight products, meant that price changes no longer required workers to remove, and re-apply, price stickers. A second consequence of this automation was that scanning stores were able to change prices more frequently.\(^5\)

### 3 Data

Data on scanner installations come from the Food Marketing Institute (FMI) publication *Scanning Installation Up-Date* and provide the month and year of installation by store, from the first installation of a National Cash Register (coincidentally, the same company that had produced the first mechanical cash register) scanner at a Marsh supermarket in Troy, Ohio, in 1974 until the end of 1984. The data were compiled by FMI through regular phone calls to scanner manufacturers, including IBM and NCR.\(^6\) These data are described in detail in Basker (2012).

Ideally, these scanner installations would be matched to high-frequency pre- and post-installation price data at the store level. Unfortunately, however, store-level price data spanning early scanner introductions do not exist, at least not on a large scale.\(^7\) Instead, I

\(^5\)The best estimates of the cost of price changes come from Levy, Bergen, Dutta, and Venable (1997), who find that price at changes cost (in 1991-1992 dollars) approximately 0.52 cents per price change in supermarkets not subject to IPLs, and $1.33 in a supermarket chain subject to such a law. What this implies is that item pricing costs as much as 81 cents per product in 1991-1992 dollars, or $1.33 in 2012 dollars. These higher costs are offset by much lower frequency of price changes, only about 40% of the frequency of non-IPL stores.

\(^6\)I thank Sue Wilkinson from FMI for describing the collection process to me.

\(^7\)Most micro-level pricing studies today use scanner data, from sources like AC Nielsen’s HomeScan database (e.g., Hausman and Leibtag, 2007), Information Resources, Inc. (e.g., Krishnamurthi and Raj, 1991), or from specific retailers that have shared store-level data, such as Dominick’s Finer Foods (e.g., Chevalier, Kashyap, and Rossi, 2003) or other retailers (e.g., Einav, Leibtag, and Nevo, 2010; Singh, Hansen, and Blattberg, 2006). Unfortunately, these types of sources are useless for the current analysis, which requires data spanning the very introductions of scanners. Store-level prices collected by the Bureau of Labor Statistics (BLS) for the purposes of calculating the Consumer Price Index (CPI) have been used by some researchers, including Cortes (2008) and Matsa (2011), but available records only go back to the mid-1980s.
use city-level average-price data for the period 1972–1984. These come from the American
Chamber of Commerce Research Association, or ACCRA (today, the Council for Community
and Economic Research, or C2ER), which started collecting prices at several hundred U.S.
cities in 1968 for the purpose of creating a “cost-of-living” index. The data are updated
quarterly, on the first week of each quarter (in January, April, July, and October). Several
stores (usually at least five and up to ten) are surveyed in each city for prices of several well-
specified products, such as “one dozen grade A eggs” or “jar of Gerber’s strained vegetables”
(baby food). For this analysis, I use ACCRA’s city-level average prices for several products
that remained consistent throughout the period 1972–1984. Stores in the ACCRA sample
are generally chain supermarkets, rather than specialists or limited-assortment grocers.

The products are listed in Table 1, by general product category. “Non-grocery” items are
items sold either exclusively (as in the case of gasoline) or largely in non-grocery stores (drug
stores, general-merchandise stores, convenience stores, and the like); these non-grocery items
should be impacted minimally, if at all, by the introduction of scanners in grocery stores.
Prices range from an average of 64 cents for a jar of Gerber’s baby food to $26.52 for a
gallon of regular gasoline, all in constant December 2012 dollars (deflated using the monthly
all-items CPI). Price series (mean and inter-quartile range) for four products with varying
degrees of cross-sectional variance and time-series volatility are shown in Figure 2.

To match the aggregation levels of the scanners and prices, I aggregate the scanner
numbers by city and create a panel of cities with the installed scanner base and prices at the
beginning of each quarter through the end of 1984. I supplement this panel with data from
the 1977 Economic Census, which provides the number of food-selling establishments (SIC

---

8 The baby-food product definition changed in 1984, switching from Gerber’s to the cheapest available; for
this reason I truncate the baby-food price series at 1983q4.

9 City-level ACCRA data have been used for a variety of economic studies in the past, including studies of
supermarket financing (Chevalier, 1995; Chevalier and Scharfstein, 1996); price convergence and deviations
from the “law of one price” (Parsley and Wei, 1996; Choi and Wu, 2012); inequality (Frankel and Gould,
2001); and the impact of retailer entry on prices (Basker, 2005; Courtemanche and Carden, 2012). To my
knowledge, only one paper has used store-level data from ACCRA (Basker and Noel, 2009); the organization
has not kept records of store-level data from the 1970s and 1980s.
54, including supermarkets and smaller grocers, specialists, and other food sellers) by city.

The sample of cities is restricted by two criteria. First, a city has to appear in the ACCRA data set in at least 20 quarters between 1974q1 and 1984q4. Second, the city has to be listed in the published 1977 Economic Census tables. This leaves 185 cities in 41 states, with number of food-selling establishments in 1977 ranging from 7 (Vermillion, SD and Holdrege, NE) to 2,809 (Chicago). Sixteen cities appear in the data in each of the 44 quarters from 1974q1 to 1984q4, and sixteen others appear in the data in just 20 quarters after 1974q1. The average number of quarters a city appears in the data after 1974q1 is 34.

Because there is considerable variation in the size of cities, I use the number of food-selling establishments in 1977 to scale the number of scanners. This scaling is necessarily imperfect, because stores opened and closed between 1972 and 1984, but it provides a ballpark.\textsuperscript{10}

Of the approximately 150,000 food-selling establishments in the U.S. in 1977, 32,000, or 21\%, were in the sample cities. Similarly, of the more than 11,000 scanners installed in U.S. grocery stores by the end of 1984, 2,487, or about 21\%, were in the sample cities. Figure 1 shows the total number of scanners installed in the U.S. over time (dashed black line, left axis) and the total number of scanners installed in the sample cities (solid red line, right axis). The earliest scanner installations in this sample of 185 cities were in Houston and Indianapolis, both in the third quarter of 1975. Fifteen cities in the sample still did not have a single scanner installation by the end of 1984. By the end of 1984, the largest number of scanners (225) was installed in Houston, and the largest number of scanner as a fraction of 1977 food-selling establishments was in Champaign, IL (13/30 \approx 0.43). City-level summary statistics are in Table 2.

\textsuperscript{10}Annual counts of stores, by type, are available at the county level from the Census Bureau’s County Business Patterns, but no equivalent exists at the city level. Zipcode Business Patterns is a recent dataset, beginning in the 1990s, and covers only a subset of U.S. zipcodes. Annual counts also risk some endogeneity, if scanners either slowed down or sped up natural churn in the grocery sector.
4 Scanners and Prices

4.1 Difference-in-Difference Estimates

To determine the average effect of scanners on city-level prices, I estimate the following difference-in-difference equation:

\[
\ln(\text{price})_{ijt} = \alpha_{ij} + \delta_{jt} + \theta \text{ScanFrac}_{it} + \varepsilon_{ijt}
\] (1)

where \( \text{price}_{ijt} \) is the average price of product \( j \) in city \( i \) in quarter \( t \), \( \alpha_{ij} \) is a city-product fixed effect, \( \delta_{jt} \) is a product-time fixed effect, and \( \text{ScanFrac}_{it} = \left( \frac{\text{Scanners}_{it}}{\text{FoodEstabs}_{i,1977}} \right) \) is the number of scanners installed in city \( i \) at the beginning of quarter \( t \), scaled by the number of food-selling establishments recorded for that city in the 1977 Economic Census. Standard errors \( \varepsilon_{ijt} \) are clustered at the city level to allow for arbitrary autocorrelation in the error term as well as arbitrary correlations in the error term across products within a city, both contemporaneously and across time periods.

The key identification assumption in Equation (1) is that prices in cities with more/earlier scanner installations and cities with fewer/later installations would have received i.i.d. shocks but for the scanners; put differently, if the Ad-Hoc Committee had not agreed on a symbol in 1973 and scanning had not been possible, no systematic differences would have been seen between the price changes in early-adopting cities and those in late-adopting cities.

There are some data-quality concerns for the price data, both because of aggregation issues and because the data are collected by employees and volunteers who do not have special training in survey methodology. Although I do not have information about the instructions given to price collectors during the period of study, in recent years ACCRA has supplied a detailed instructions manual to price collectors, including the types of stores they should survey (stores where professionals shop), the number of stores (at least five), what to do in the event the product is out of stock, etc. Perhaps more importantly, because prices are
a left-hand side variable in this analysis, any i.i.d. errors should be incorporated into $\varepsilon_{ijt}$, and any systematic bias at the product, city, or product × city level will be captured by the fixed effects $\alpha_{ij}$. The only remaining concern, then, is that systematic errors or biases in price collection were correlated with the introduction or diffusion of scanners; this concern is addressed in several specification tests in Section 4.3.

Pooling all observations into a single regression, the point estimate for $\theta$ in the pooled regression is approximately $-0.15$, statistically significant at conventional levels. This figure is shown in the first column of Table 3. To interpret this coefficient, note that the median value of ScanFrac in the data, for city-quarter observations in which at least one store has already installed a scanner, is 5%. Extrapolating the effect of scanners to the case where ScanFrac approaches one is far out of sample. Many of the establishments included in SIC 54 are not the sort of large grocery stores and supermarkets that installed scanners in the early years; in the 1977 Census of Retail Trade, 30% of SIC 54 stores were specialists including meat and fish markets and bakeries, and another unknown fraction were convenience stores and other small grocers. Instead, I evaluate the estimates at ScanFrac = 5%. The product $0.05 \cdot \hat{\theta}$ provides an estimate of the short-run impact of the fraction of scanning grocery stores in the city within the range for which data are available: an average price decrease of 0.75%.\textsuperscript{11}

To get a sense of the magnitude of these figures, if demand at an individual supermarket has a constant elasticity, then price is a constant markup over marginal cost, $p = \frac{\varepsilon}{\varepsilon - 1} c$, and the marginal-cost elasticity of price $\left( \frac{dp}{dc} \right)_{p}$ is 1, so a 0.75% decrease in price implies a 0.75% decrease in marginal cost for the average store. For stores that installed scanners, Basker (2012) finds that scanners reduced store-level labor costs by 4.5% on average. Industry-level estimates of the fraction of costs accounted for by labor costs are not available, but the 1979

\textsuperscript{11}This discussion implies a nonlinear effect of scanner penetration. Estimating a nonlinear effect by adding a squared term to the regression equation does produce the expected signs, but the power of the regression is very low, and neither the linear nor the squared term is statistically significant.
annual report for Kroger (a chain with 1,234 stores at the time) indicated that operating, general, and administrative expenses, which include mostly labor costs, constituted 18.3% of costs that year. If this number is representative of the industry, and if the 4.5% decrease in total labor costs found in Basker (2012) also translates into a 4.5% decrease in marginal labor cost, then marginal cost fell by $0.045 \cdot 0.183 \approx 0.82\%$, consistent with the price effect I estimate here. This calculation assumes that the decrease in prices is uniform for both adopting and non-adopting stores, perhaps due to competitive pressures from adopting stores. Alternatively, marginal cost at adopting stores may have fallen more than average cost, and non-adopting stores may have adjusted their prices only partially.

To better understand the mechanism for this effect, I next test for heterogeneous price effects over two margins: cities and products.

### 4.2 Heterogeneity

#### 4.2.1 Product-Level Heterogeneity

To test whether the price reductions were uniform across products, I re-estimate Equation (1) separately by product category. These results are shown in columns 2-6 of Table 3. Each pooled regression includes three products, by category — produce, dairy, meat, canned, and miscellaneous grocery items. As above, I allow both the time fixed effects and the city fixed effects to vary by product, and cluster the standard errors at the city level to allow for arbitrary correlations in the error term both within and across products. The coefficients are quite dramatically different across product category, with the largest price decreases — corresponding to approximately 1.5% and 1.1% at the 5% adoption levels — for produce and meat products, smaller (and statistically insignificant) decreases for dairy and miscellaneous items, and no price response for canned goods. Repeating this analysis product-by-product, I find that most of the variation in $\hat{\theta}$ occurs across, rather than within, product categories. To conserve space, the predicted price effects for each of the 15 products are reported in Figure 3. Fourteen of the 15 effects are negative; of these, three (bananas, lettuce, and
chicken) are significant at the 1% level, and three more (potatoes, bacon, and shortening) are significant at the 5% level. The single positive coefficient, not statistically significant, is for the price of canned tomatoes.

The exact dates at which individual products started featuring barcodes are not available, but we do know that several prominent canned-goods producers were early members of the UGPCC. Hunt-Wesson, manufacturer of Hunt’s canned tomatoes, had three separate UGPCC registrations as of July 1974, as did Del Monte, another large manufacturer of canned goods; Nestlé (which had acquired Libby’s, a canned-food manufacturer, in 1971) had four registrations (Distribution Codes, Inc., 1974). As noted earlier, each registration allowed a company to print up to 100,000 separate UPC symbols, so three registrations would have covered as many as 300,000 products. It seems likely that for most, if not all, the period of this study, supermarket scanners would have been able to scan canned tomatoes and similar products.

The fact that produce and meats, not canned goods, exhibited the largest price reductions is consistent with the observation that scanners were often coupled with other technologies, specifically electronic cash registers, which allowed the use of price look-up codes. These codes saved cashiers the need to memorize an ever-changing list of prices or waste valuable time looking up these prices.

4.2.2 Spatial Heterogeneity

When scanning was still in its infancy, several states — under pressure from unions and consumer groups — passed so-called “item-pricing” laws (IPLs), requiring stores to retain item pricing even in the event that technological advancement made this practice obsolete. The Massachusetts Attorney General passed an item-pricing regulation as early as 1971, with the goal “to ensure that check-out clerks were not attempting to charge prices by memory, and to prevent pricing discrimination between customers” (Hurst, 2007). As of February 1976, three states – California, Connecticut, and Rhode Island — had passed such laws, and
eight more were considering such action (Wilson, 1976); by the end of 1976, Michigan and New York had also passed IPLs (Das, Falaris, and Mulligan, 2009). Several of these laws are still in place today.\textsuperscript{12}

To test whether item-pricing laws limited the cost-saving (and price pass-through) of scanners, I estimate Equation (1) separately on the 165 cities in states with no IPLs and on the 20 cities in my sample in states with IPLs. The results are given in Table 4, with the top panel reproducing Table 3 for non-IPL cities, and the bottom panel for IPL cities. The estimated effect of scanners in non-IPL cities is qualitatively similar, and slightly larger, than the effect in the full sample, with standard errors generally only slightly larger than in the full sample. (An exception is the estimate for the effect of scanners on prices of dairy products; there the coefficient increases in absolutely value and the standard error decreases so that the point estimate becomes significantly different from zero at the 10% level.) The IPL sample, however, is too small — particularly with clustering at the city level — to provide any statistically meaningful results. Standard errors large: in the case of the full set of products, the standard error is two orders of magnitude greater than the coefficient. While these results are far from conclusive, they are consistent with larger effects of scanners in cities with no item-pricing requirements.

More generally, we may expect that scanners had more potential to lower costs in locations with a lower union presence than in those with highly organized grocery workers. I test this hypothesis by estimating a variant of Equation (1) in which I introduce a unionization variable, which I interact with both the scanner variable and the time×product fixed effects.\textsuperscript{13} Across specifications, neither the coefficient on the scanner variable nor the coeffi-

\textsuperscript{12}Levy, Bergen, Dutta, and Venable (1997), Bergen, Levy, Ray, Rubin, and Zeliger (2008), and Levy (2008) make a compelling case that IPLs are costly to stores and, indirectly, to consumers. Nakamura (1999) and others also note that the eventual elimination of item prices increased the ease of price changes, allowing retailers to experiment and learn about price elasticities.

\textsuperscript{13}Unionization data are not available for the entire historical period covered in this paper; the Current Population Survey (CPS) started asking about union membership in 1983. Nor are the data available at the city level, but the CPS does have state identifiers. To calculate the relevant unionization rate I combined data from the CPS Outgoing Rotation Groups (ORG) for 1983-1986, keeping only respondents whose union
cient on the interaction term $\text{ScanFrac}_{it} \times \text{Union}_i$ is ever statistically significant, possibly due to significant measurement error in the union variable (which is constructed at the state level from a fairly small sample and over a period extending past the end of the scanner sample). Absent finer and earlier data on grocery-sector unionization, it is impossible to draw any conclusions from these regressions. These results are available upon request.

4.3 Other Explanations

The identification of the effects of scanners on prices above depends on the standard difference-in-difference assumption that the timing of treatment (in this case, scanner adoption) is independent of the error term. While I cannot test this assumption directly, I perform two “placebo tests” for endogeneity/omitted-variable bias in this section. I report the results of these exercises in the next two subsections, and follow up with a discussion.

4.3.1 Lead Effects

One omitted-variable concern is that stores may have installed scanners because prices were anticipated to decline for other reasons. To test for this possibility, Table 5 shows the results of an augmented regression in which next quarter’s scanner adoption rate, $\text{ScanFrac}_{i,t+1}$, is included along with the current quarter’s:

$$\ln(price)_{ijt} = \alpha_{ij} + \delta_{jt} + \beta \text{ScanFrac}_{i,t+1} + \theta \text{ScanFrac}_{it} + \varepsilon_{ijt}$$ (2)

The coefficient on the lead of the scanner adoption rate, $\beta$, tells us whether prices started changing immediately prior to an increase in the scanner penetration rate in a city. A signif-
icant coefficient on the lead scanner variable could indicate reverse causality of the estimates in Table 3. A zero coefficient, in contrast, lends credence to the exogeneity assumption.

Although the standard errors are large, there is no evidence that price decreases preceded or anticipated scanner diffusion; if anything, the opposite is true. The effect on the lead coefficient is positive for regression with all products and for four of the five product-category regressions; it is negative, but extremely small, for the fifth. The estimated price effect of current-period scanning increases (in absolute value) for all product categories, in contrast, becoming significant at the 5% level for dairy products and at the 10% level for miscellaneous products, despite much larger confidence intervals.

4.3.2 Non-Grocery Products

In addition to prices of grocery products, ACCRA collects and reports prices of several non-grocery items. These include items that are rarely sold in grocery stores, like gasoline, and items sold by both grocery stores and many non-grocery (and hence, in this time period, non-scanning) stores including general merchandisers and drug stores, such as cigarettes. I re-estimate Equation (1) using three of these products, whose definitions are constant over time: cigarettes, gasoline, and alcohol. According to the 1977 Census of Retail Trade, grocery stores accounted for 42% of all alcohol sales by value, 23% of all tobacco sales, and less than 2% of all automotive fuel and lubricant sales. (By contrast, grocery stores accounted for 89–93% of sales of dairy products, fresh produce, and meat.) These prices may be minimally affected by scanning, but the effects should be much smaller than that for grocery products.

Estimation results are shown in Table 6. For the three non-grocery products combined, the coefficient on the scanner variable is very small and tightly estimated. Of the three individual products, two have coefficients and one has a positive coefficient, though none are statistically significant. The largest negative effect is on liquor prices; this is the product whose share of sales in grocery stores is the largest of the three, so we may expect some causal effect of scanning on this price.
Overall, the results suggest that there are no overall declining price trends that can provide an alternative explanation for the main results in the paper.

This result is not conclusive for at least two reasons. First, the non-grocery products are significantly more expensive than the grocery products in the main analysis, so they are not strictly comparable on this important dimension. Second, alcohol, cigarettes, and gasoline are products often heavily regulated, and in particular may be subject to resale-price maintenance regulation in some jurisdictions, in which case their prices may not reflect competitive pressures. Nevertheless, the results are consistent with the hypothesis that scanner adoption, and not some omitted variable, is driving grocery-price decreases.

4.3.3 Interpretation

There are other possible reasons for the observed price declines that are not attributable to lower labor costs at checkout.

One possibility is that at least some of the decline in prices is due to product changes. While the product definitions in Table 1 remained constant over the study period, most of them did not specify brands. If scanner adoption facilitate the substitution of cheaper for more expensive brands, or of store brands (generics, or “private-label” products) for branded ones, the effect may be a brand effect rather than a labor-cost effect. McEnally and Hawes (1984) and Ward, Shimshack, Perloff, and Harris (2002) point to the late 1970s as the watershed period for the introduction of store-brand groceries. Later scanner-generated data shows that private labels’ market increased across product categories in the late 1980s and early 1990s (Hoch, Montgomery, and Park, 2000), but comparable data spanning the introduction of scanners is lacking.

If store-brand introduction were uniform across the sample cities, the time fixed effects in the regression equations should absorb this effect. If store brands sometimes preceded, and sometimes followed, scanners, we should have seen negative, not weakly positive, lead effects in Section 4.3.1. If store brands systematically coincided with scanners, we should
have seen an impact on scanners in both non-IPL and IPL cities, in Section 4.2.2. In other words, if the estimated effect of scanning is driven by product substitution, the scanners would have had to have coincided with (but not preceded by) store brands, but only in non-IPL cities. This could have happened if scanners, because of the labor-cost reduction they induced where item-prices were eliminated, allowed supermarkets to expand their product selection by simplifying inventory management, or though a more complicated mechanism as in Holmes (2001) or Basker, Klimek, and Van (2012). As more products and more varieties of each product are introduced, the price distribution within each product type (e.g., milk) increases, and the first order statistic (minimum) of the price distribution within each store weakly decreases. ACCRA prices for many products are given as the average, across stores, of the lowest price within each store.

This explanation is hard to test in the current data. Only two of the grocery products in the ACCRA dataset are branded over this time frame: baby food (Gerber) and shortening (Crisco). While the effect of scanners on the price of baby food is nil (Figure 3), the price effect for shortening is well within the range of non-branded products.

Another possibility is that the price changes are due to demand factors. Although it is plausible that consumers shifted away from scanning stores at first, due to distaste for the technology, demand factors seem unlikely to be able to drive these results because the prices here are average (city-level) prices and aggregate demand for groceries is extremely inelastic. Consumers who shifted away from scanning stores (leading those stores to lower prices) would presumably be shifting towards non-scanning stores, leading these to raise prices; average prices are unlikely to have changed much as a result of such a demand realignment.

The most plausible explanation, then, is marginal-cost pass-through.
5 Conclusion and Welfare Implications

The above analysis implies that a conservative estimate of the full impact of scanners — before later add-on innovations such as inventory management, electronic payments, and the like, but in conjunction with complementary technologies of the day, specifically price look-up codes — is a decrease in U.S. grocery prices by about 0.75% in the short run. This estimate is conservative in that it assumes nonlinear price effects so that the full effect is realized during the sample period (by the mid-1980s), with an adoption rate of 5% on average, and that it does not incorporate either later innovations or the effect of scanners on prices in other sectors, such as apparel, that adopted the technology later.

The demand for food is quite inelastic, especially outside of the very poorest populations, so a 0.75% decrease in the price of food is associated with an increase in consumer surplus amounting to approximately 0.75% of total food expenditure. The 1977 Census of Retail Trade calculates consumers' total expenditure on groceries and other food (merchandise line 100) at $137.5 billion in 1977 dollars, or nearly $400 billion in 2012 dollars. The 1982 Census of Retail Trade calculates retail expenditures on merchandise line 100 at nearly $200 billion in current dollars, or $470 billion in 2012 dollars. An increase in consumer surplus equivalent to 0.75% of food expenditure is therefore on the order of $3.25 billion in 2012 dollars.

This effect is large compared to other consumer-surplus increasing interventions. Whitmore Schanzenbach (2002), for example, estimates the consumer-surplus gain from replacing food stamps with cash-in-kind payments at about $500 million in 2002 dollars, or about $640 million in 2012 dollars, about a fifth of the benefit I estimate from early scanners.

This calculation is a lower bound in several ways. First, the ACCRA prices used here do not account for substitution across stores or products as a result of changes in relative prices. Stores in the ACCRA sample are weighted equally, so the only substitution these prices allow are due to store entry/exit from the sample. This, in turn, can occur non-randomly if stores open or shut down or if stores become less attractive and the surveyor determines that professionals are no longer likely to shop there. While these effects do introduce some degree
of substitution into the data, the average quantity-weighted price is likely to be much lower than the average price reported by ACCRA. Whether the quantity-weighted price declines more or less sharply depends on which stores lower their prices and on the elasticity of substitution in the demand function.\textsuperscript{14}

Second, barcodes were just the first domino in a long chain of innovations that have transformed the food supply chain, and over time other sectors as well. The current estimates pertain only to the very earliest savings, due to faster checkout times and lower costs of price changes, and do not incorporate improvements the occurred after 1984. To the extent that barcodes really did lead to frequent deliveries and superstores (Holmes, 2001), they have helped chains generate economies of scale and scope and have had an even larger impact on prices by creating opportunities for the reorganization of the supply chain. Quantifying these long-run effects is considerably more challenging due to the intervention of confounding factors, but the effects estimated in this paper suggest that barcodes induced large effects even in the short run, before these add-on complementary technologies were introduced.

The effect of scanning on consumer prices is ironic in historical context. Resistance to scanning was strong among consumer groups, which (correctly) believed scanning to be a precursor to item-pricing removal. Before the UPC, every price change required workers to remove, then re-apply, a price sticker on every item in stock. Unions, worried that grocery-workers’ hours would be cut if item prices were eliminated, persuaded consumers that they, too, would suffer from such a change by losing the ability to quickly and transparently compare prices across products and stores and to verify that they were not overcharged at checkout. Consumers were also concerned, in the high-inflation 1970s, that item-price removal would lead supermarkets to accelerate the frequency of price increases. As it turns out, consumers benefited from the new technology.

\textsuperscript{14}In a different setting altogether, Basker and Noel (2009) found 1\% average-price reductions at incumbent supermarkets when Wal-Mart Supercenters opened in town, while Hausman and Leibtag (2007) find a quantity-weighted effect that is four times larger.
References


Figure 1. U.S. and Sample Scanner Installations, 1972–1985q1

Figure 2. Mean and Inter-Quartile Range of Six Real Price Series
Figure 3. Estimated SR Effects of 5% Scanner Penetration, by Product
Table 1. Retail Price Summary Statistics

<table>
<thead>
<tr>
<th>Product</th>
<th>Description</th>
<th>Obs</th>
<th>Average Price</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Produce:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bananas</td>
<td>1 lb.</td>
<td>7,058</td>
<td>0.91</td>
<td>0.17</td>
</tr>
<tr>
<td>Lettuce</td>
<td>Head iceberg lettuce, approx. 1½ lbs.</td>
<td>7,059</td>
<td>1.75</td>
<td>0.45</td>
</tr>
<tr>
<td>Potatoes</td>
<td>10 lbs., red or white</td>
<td>7,029</td>
<td>4.79</td>
<td>1.61</td>
</tr>
<tr>
<td><strong>Dairy:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eggs</td>
<td>1 dozen, grade A, large</td>
<td>7,058</td>
<td>2.58</td>
<td>0.62</td>
</tr>
<tr>
<td>Milk</td>
<td>½ gallon whole milk, carton, fresh</td>
<td>7,058</td>
<td>3.50</td>
<td>0.44</td>
</tr>
<tr>
<td>Margarine</td>
<td>1 lb., cheapest</td>
<td>7,058</td>
<td>1.44</td>
<td>0.42</td>
</tr>
<tr>
<td><strong>Meat:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beef</td>
<td>Ground beef, 1 lb.</td>
<td>7,059</td>
<td>3.71</td>
<td>0.78</td>
</tr>
<tr>
<td>Bacon</td>
<td>1 lb., cheapest</td>
<td>7,058</td>
<td>4.27</td>
<td>1.70</td>
</tr>
<tr>
<td>Chicken</td>
<td>Frying chicken, grade A, whole, per lb.</td>
<td>7,059</td>
<td>2.10</td>
<td>0.44</td>
</tr>
<tr>
<td><strong>Canned:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peaches</td>
<td>#2½ (approx 29 oz.) can halved peaches, cheapest</td>
<td>7,059</td>
<td>2.30</td>
<td>0.26</td>
</tr>
<tr>
<td>Baby food</td>
<td>4.5 oz. jar Gerber’s strained vegetables</td>
<td>6,256</td>
<td>0.64</td>
<td>0.80</td>
</tr>
<tr>
<td>Tomatoes</td>
<td>#303 can (15–17 oz.), cheapest</td>
<td>7,059</td>
<td>1.25</td>
<td>0.19</td>
</tr>
<tr>
<td><strong>Misc.:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orange Juice</td>
<td>6 oz. can, frozen, cheapest</td>
<td>7,055</td>
<td>1.17</td>
<td>0.18</td>
</tr>
<tr>
<td>Bread</td>
<td>20 oz. white bread, cheapest</td>
<td>7,058</td>
<td>1.55</td>
<td>0.32</td>
</tr>
<tr>
<td>Shortening</td>
<td>3 lbs. can all vegetable shortening, Crisco brand</td>
<td>7,058</td>
<td>6.19</td>
<td>1.50</td>
</tr>
<tr>
<td><strong>Non-Grocery:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gasoline</td>
<td>1 gallon, unleaded regular grade, including</td>
<td>7,057</td>
<td>26.52</td>
<td>5.17</td>
</tr>
<tr>
<td>Liquor</td>
<td>taxes, national brand, self-service if available</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cigarettes</td>
<td>Seagram’s 7 Crown, 750 ml.</td>
<td>6,989</td>
<td>19.55</td>
<td>4.23</td>
</tr>
<tr>
<td></td>
<td>Carton, Winston, king size</td>
<td>7,058</td>
<td>16.85</td>
<td>2.26</td>
</tr>
</tbody>
</table>

All prices are in December 2012 dollars
Most missing observations are due to unreadable hard-copy
a Gerber’s baby food is available only to 1983q4
Table 2. City Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average</th>
<th>Standard Deviation</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977 Food-Selling Establishments</td>
<td>174.3</td>
<td>293.8</td>
<td>7</td>
<td>2,809</td>
</tr>
<tr>
<td>Scanners</td>
<td>3.3</td>
<td>10.6</td>
<td>0</td>
<td>231</td>
</tr>
<tr>
<td>100×Scanners/Estabs</td>
<td>2.2</td>
<td>4.5</td>
<td>0</td>
<td>43.3</td>
</tr>
</tbody>
</table>

Table 3. Retail Prices as a Function of Scanners

<table>
<thead>
<tr>
<th>ScanFrac</th>
<th>All</th>
<th>Produce</th>
<th>Dairy</th>
<th>Meat</th>
<th>Canned</th>
<th>Misc.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-0.1499***</td>
<td>-0.3108***</td>
<td>-0.0625</td>
<td>-0.2302***</td>
<td>-0.0015</td>
<td>-0.1295</td>
</tr>
<tr>
<td></td>
<td>(0.0499)</td>
<td>(0.0845)</td>
<td>(0.0608)</td>
<td>(0.0643)</td>
<td>(0.0529)</td>
<td>(0.0818)</td>
</tr>
<tr>
<td>Products</td>
<td>15</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Product×City FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Product×Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>105,041</td>
<td>21,146</td>
<td>21,174</td>
<td>21,176</td>
<td>20,374</td>
<td>21,171</td>
</tr>
<tr>
<td>SR Effect</td>
<td>-0.0075</td>
<td>-0.0155</td>
<td>-0.0031</td>
<td>-0.0115</td>
<td>-0.0001</td>
<td>-0.0065</td>
</tr>
</tbody>
</table>

Each column represents a separate regression
Robust standard errors in parentheses, clustered by city

* p<10%; ** p<5%; *** p<1%

The estimated SR effect is 0.05·, where is the coefficient on ScanFrac
Table 4. Retail Prices as a Function of Scanners: IPL vs. non-IPL

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Produce</th>
<th>Dairy</th>
<th>Meat</th>
<th>Canned</th>
<th>Misc.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Non-IPL Cities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ScanFrac</td>
<td>-0.1646***</td>
<td>-0.3361***</td>
<td>-0.0934*</td>
<td>-0.2431***</td>
<td>0.0038</td>
<td>-0.1376</td>
</tr>
<tr>
<td></td>
<td>(0.0510)</td>
<td>(0.0865)</td>
<td>(0.0551)</td>
<td>(0.0672)</td>
<td>(0.0549)</td>
<td>(0.0891)</td>
</tr>
<tr>
<td>Products</td>
<td>15</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Product×City FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Product×Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>92,489</td>
<td>18,624</td>
<td>18,651</td>
<td>18,653</td>
<td>17,912</td>
<td>18,649</td>
</tr>
<tr>
<td>SR Effect</td>
<td>-0.0082</td>
<td>-0.0168</td>
<td>-0.0047</td>
<td>-0.0122</td>
<td>0.0002</td>
<td>-0.0069</td>
</tr>
</tbody>
</table>

| **IPL Cities** |       |         |       |        |        |        |
| ScanFrac        | -0.0037 | 0.0404   | 0.2141 | -0.0650 | -0.1538 | -0.0721 |
|                 | (0.1468) | (0.3547) | (0.1892) | (0.1579) | (0.1510) | (0.1635) |
| Products        | 15    | 3       | 3     | 3      | 3      | 3      |
| Product×City FE | ✓     | ✓       | ✓     | ✓      | ✓      | ✓      |
| Product×Time FE | ✓     | ✓       | ✓     | ✓      | ✓      | ✓      |
| Observations    | 12,552 | 2,522    | 2,523 | 2,523  | 2,462  | 2,522  |

Each column represents a separate regression
Robust standard errors in parentheses, clustered by city
* p<10%; ** p<5%; *** p<1%

The estimated SR effect is $0.05 \cdot \hat{\theta}$, where $\theta$ is the coefficient on ScanFrac

Table 5. Retail Prices as a Function of Current and Future Scanners

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Produce</th>
<th>Dairy</th>
<th>Meat</th>
<th>Canned</th>
<th>Misc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ScanFrac$_{i,t+1}$</td>
<td>0.1069</td>
<td>-0.0021</td>
<td>0.1511</td>
<td>0.0962</td>
<td>0.1048</td>
<td>0.1804</td>
</tr>
<tr>
<td></td>
<td>(0.0946)</td>
<td>(0.1548)</td>
<td>(0.1035)</td>
<td>(0.1396)</td>
<td>(0.1174)</td>
<td>(0.1421)</td>
</tr>
<tr>
<td>ScanFrac</td>
<td>-0.2559***</td>
<td>-0.3088**</td>
<td>-0.2123**</td>
<td>-0.3256**</td>
<td>-0.1056</td>
<td>-0.3083*</td>
</tr>
<tr>
<td></td>
<td>(0.0962)</td>
<td>(0.1422)</td>
<td>(0.0995)</td>
<td>(0.1442)</td>
<td>(0.1256)</td>
<td>(0.1656)</td>
</tr>
<tr>
<td>Products</td>
<td>15</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Product×City FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Product×Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>105,041</td>
<td>21,146</td>
<td>21,174</td>
<td>21,176</td>
<td>20,374</td>
<td>21,171</td>
</tr>
</tbody>
</table>

Each column represents a separate regression
Robust standard errors in parentheses, clustered by city
* p<10%; ** p<5%; *** p<1%
Table 6. Non-Grocery Retail Prices as a Function of Scanners

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Gasoline</th>
<th>Liquor</th>
<th>Cigarettes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ScanFrac</strong></td>
<td>-0.0147</td>
<td>-0.0389</td>
<td>-0.0786</td>
<td>0.0732</td>
</tr>
<tr>
<td></td>
<td>(0.0301)</td>
<td>(0.0407)</td>
<td>(0.0804)</td>
<td>(0.0580)</td>
</tr>
<tr>
<td>Products</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Product×City FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Product×Time FE</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Observations</td>
<td>21,104</td>
<td>7,058</td>
<td>7,057</td>
<td>6,989</td>
</tr>
</tbody>
</table>

Each column represents a separate regression
Robust standard errors in parentheses, clustered by city
* p<10%; ** p<5%; *** p<1%