Raising the Barcode Scanner: Technology and Productivity in the Retail Sector

Emek Basker*
University of Missouri

September 2011

Abstract

Barcodes and barcode scanners transformed the grocery industry in the 1970s. I use store-level data from the 1972, 1977, and 1982 Census of Retail Trade, matched to data on store scanner installations, to estimate scanners' effect on labor productivity. I find that early scanners increased a store's labor productivity, on average, by approximately 4.5 percent in the first few years. The effect was larger in stores carrying more packaged products, consistent with the presence of network externalities. Short-run gains were small relative to fixed costs, suggesting that the impediment to widespread adoption of the new technology was profitability, not coordination problems.

JEL Codes: L81, D22, O33

Keywords: Barcode scanners, Retail, Supermarkets, Technology, Productivity

*Comments welcome to emek@missouri.edu. This paper was written while I was an ASA/NSF/Census Bureau Fellow visiting the Center for Economic Studies (CES) at the U.S. Census Bureau. I thank the funding agencies for their generous support and the economists at CES for their hospitality. Any opinions and conclusions expressed herein are those of the author and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. I thank John Meisel and Sue Wilkinson for help obtaining the Food Marketing Institute reports. Thanks also to Saku Aura, Roger Betancourt, Teresa Fort, Lucia Foster, Shane Greenstein, John Haltiwanger, Shawn Klimek, Mark Lewis, David Matsa, Guy Michaels, Javier Miranda, Jim Mulligan, Justin Pierce, Tim Simcoe, David Weil, and seminar participants at U.S. Census Bureau, Bureau of Economic Analysis, Bureau of Labor Statistics, UNC-Greensboro, the 2011 IIOC (Boston) and the 2011 NBER Standards, Patents and Innovation pre-conference for helpful comments and conversations. All remaining errors and omissions are my own.

1 Introduction

This paper takes a first step towards achieving a better understanding of the effect of technology on the retail sector by investigating the productivity impact of barcode scanners, which were first installed in the mid-1970s, on productivity in supermarkets. After decades in the "idea" (pre-commercial) phase, once scanners commercialized, they caught on quickly: about a third of U.S. supermarkets adopted scanners within ten years of the first commercial installation. Scanners also provide a discrete, easy-to-measure form of technological innovation, providing a clean laboratory to study their impact on productivity.

Productivity in the retail sector is notoriously hard to measure and not well understood. The link between chains and productivity is made most convincingly by Foster, Haltiwanger, and Krizan (2006), who find that virtually all productivity growth in retailing in the 1990s was due to store entry (particularly through chain expansion) and exit (particularly of non-chain stores). This finding is consistent with a technology gap between chains and "Mom and Pop" retailers, although data limitations prevent the authors from determining whether this productivity gap is due to technology, management, economies of scale, or some other factor. Doms, Jarmin, and Klimek (2004) come the closest to linking retail productivity and technology adoption by showing that retail firms reporting high levels of IT investment in 1992 also experienced high labor productivity growth between 1992 and 1997. In this paper, I go even further by locating a "smoking gun" (more accurately, a smoking scanner) in IT investment and linking it directly to store-level productivity growth. Because the scanners diffused over time, I use the panel dimension of the data to rule out some omitted-variable and endogeneity concerns.

By matching data from the Food Marketing Institute on store scanner installations in the 1970s and early 1980s to Census Bureau records at the establishment (store) level, I trace the impact of scanners on store-level productivity. I use data from the Census of Retail Trade for 1972, 1977, and 1982 in a difference-in-difference specification and find that productivity increased by an average of 4.5% in stores that installed barcode scanners before

1982. This average effect masks variation both over time and across stores: gains in the year of installation were smaller, and varied even in later years with the products sold by each store. To address endogeneity bias, I use only stores that installed scanners by the end of 1984. Specification tests confirm that this sample is free of sample-selection and omitted-variable bias.

I use my estimates to perform two back-of-the-envelope calculations. In the first, I provide a ballpark estimate of the increase in the productivity boost due to scanners from having one more barcoded product on the store's shelves. This increase turns out to be the equivalent of about $28 per year, per store, giving some sense of the magnitude of network externalities in the implementation of scanning. A second calculation totals up the costs, pecuniary and non-pecuniary, and benefits from scanning, and concludes that the early scanners probably did not provide a positive return on investment. This finding suggests that limited short-run profitability, not coordination costs, was the salient barrier to early adoption of this new technology.

The large gap between the short-run private returns and the long-term social return on the investment in scanners can be explained by the emergence over time of complementary technologies that would not have been possible in the absence of scanners. In the short run, the limited number of products bearing UPC symbols was a binding constraint on the potential for productivity gains. At the time of the first scanner installations, only 2,000 food and beverage manufacturers had signed up to print barcodes on their products; by the end of 1984, nearly 13,000 manufacturers had signed up, greatly increasing the value of scanners (Dunlop and Rivkin, 1997). Longer-run developments included inventory management, which, in turn, made possible electronic data interchange (EDI) between retailers and their suppliers, including electronic payments (Abernathy, Dunlop, Hammond, and Weil, 1999).

This paper builds on a small literature on scanner installations. Levin, Levin, and Meisel (1985, 1987, 1992) and Das, Falaris, and Mulligan (2009) have used different subsets of the scanner installation data to study, respectively, the reasons for variable early diffusion rates

across U.S. metropolitan areas, the role of market concentration and firm size in scanner adoption, differences in the speed of intra-firm diffusion across supermarket chains, and changes in the diffusion process with the emergence of new vintages of the technology. Beck, Grajek, and Wey (forthcoming) use aggregate data on scanner adoption in ten European countries to estimate country-wise diffusion parameters and relate them to the structure of the retail market in each country. I depart from this literature on scanner installations in two important regards. First, my focus is on the effects of scanners rather than on their diffusion pattern. Second, merging the scanner data with store-level information from the Census of Retail Trade allows me to construct a control group for adopting stores at each point in time, and compare their outcomes with a counterfactual.

The paper also contributes to the broader discourse on the impact of technology on productivity (for a flavor of this literature, see Brynjolfsson and Hitt, 2003; Jorgenson, Ho, and Stiroh, 2008; Marrano, Haskel, and Wallis, 2009). Unlike most of that literature, which has focused on the 1990s, my study uses data from an earlier era of rapid technological innovation, and from a sector — retailing — that has been largely under-studied. A final twist in this case study is the presence of two-sided network externalities. Scanner installation can only increase a supermarket's productivity if the products on its shelves are equipped with UPC symbols (barcodes). Although I do not have data on the food-manufacturer side of the market, my calculations suggest that network externalities were substantial.

The rest of the paper is organized as follows. Section 2 provides background on the grocery sector and on the history of barcode scanners. Section 3 describes the data sources used in this study. Sections 4–6 attempt to answer the following questions: What was the impact of scanners on store-level productivity? How was this productivity gain achieved? And did scanners' benefits justify their costs? Section 7 Concludes.

2 Background

The barcode, or Uniform 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. 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 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).

Scanning became a reality in June 1974 at a Marsh supermarket in Troy, Ohio.[1] Over the next ten years, the publication *Scanning Installation Up-Date* by the Food Marketing Institute (FMI) regularly reported the number of grocery stores that installed or upgraded their scanners. At first, scanner installation was slow, but it picked up in the early 1980s. In January 1985 the publication *Marketing News* reported that 29% of supermarkets in the U.S. were using the technology (reported in Das, Falaris, and Mulligan, 2009). Figure 2 shows the timeline of scanner installations based on the FMI data; by the end of 1984, more than 10,000 supermarkets had installed scanners. The shaded regions show the periods on which the present study focuses.

Scanners improved over time. The main improvements in scanning technology before the mid-1980s were, first, an increase in scanners' ability to read small bar codes, and second, the introduction of a holographic technique that allowed optical scanning of damaged or wet

[1] Coincidentally, 1974 is the year that Greenwood and Yorukoglu (1997) cite as a watershed, the year in which the pace of technological change in the efficiency of new equipment increased.

barcodes (Das, Falaris, and Mulligan, 2009). These innovations combined with a general trend of increasing UPC adoption by food manufacturers resulted in better UPC coverage across the food sector. Bill Selmeier, who worked for IBM during the 1970s, reports that when scanners were first installed, "very few items [...] had the symbol included in their brand packaging," limiting the scanners' benefits (Selmeier, 2008, p. 223). As time went on, however, packages were increasingly likely to include barcodes. According to one account, by 1980 more than 90% of grocery products had UPC registrations (Harmon and Adams, 1984, p. 7). Zimmerman (1999) reports that between May 1974 and November 1980, the number of grocery manufacturers that had adopted UPCs increased almost tenfold, from 869 to 7,570 (p. 45).

The technology was originally intended to speed up customer check-out and reduce labor costs at the cash register as well as on the store floor, for example for price changes. The main reason early adopters gave for implementing the new technology was their interest in assessing "the labor-saving potential of the equipment" in a realistic setting (Shaw, 1977, p. 54). Few initial adopters conducted proper assessments of the technology, either because they had no baseline against which to measure productivity or because they implemented other changes in conjunction with scanner installation (Shaw, 1977, p. 51). Instead, they relied on McKinsey's forecasts and the equipment manufacturers' estimates to justify the expense.[2]

Early results of scanner installations were mixed if not disappointing. After interviewing 50 retailers that adopted scanners by November 30, 1976, Shaw (1977, p. 233) concludes, "23 firms ... [reported] improved speed of throughput due to scanning at the checkout, while 12 ... [claimed] unchanged or reduced productivity. The results achieved by the remaining 15 were indeterminate." There are some reports from the late 1970s and early 1980s on the productivity gain from these scanners, but none are convincingly documented. Harmon and

[2]Brown (1997, p. 44) notes that McKinsey "was far from a disinterested, neutral observer in the process. It saw its role as that of a leader and advocate [for scanning], and was not shy about exercising it."

Adams (1984, p. 204), for example, report that the grocery industry realized productivity gains of 40% over manual price entry, but cite no source for this figure.

Today, scanners provide previously unimaginable data. Store managers report using scanner data primarily for promotions and price setting (Bucklin and Gupta, 1999), but scanners can also provide detailed worker-level productivity information — items scanned per second — which may be used in promotion or compensation decisions. Mas and Moretti (2009) use such detailed productivity data to study the effect of peers on checkout speed. Scanner data also helps track consumer demand through so-called "loyalty" cards, and scanner-enabled inventory management has improved in-stock rates (Matsa, forthcoming). Many authors have speculated that barcode scanners — and the IT revolution in inventory management that scanning made possible — provided the foundation for the increased product selection and the growth of stores we have observed in recent decades (see, e.g., Holmes, 2001; Basker, Klimek, and Van, forthcoming).

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 scanner installation at a Marsh supermarket store in Troy, Ohio, in June 1974, until the end of 1984. The data were compiled by FMI through regular phone calls to scanner manufacturers, including IBM and National Cash Register (NCR).[3] More than 12,000 installations are listed in these files by the end of 1984, approximately a quarter of which are specified to be "upgrades." I focus my attention on *de novo* installations. Figure 3

[3]I thank Sue Wilkinson from FMI for describing the collection process to me. FMI continued to publish reports through December 1985, but later reports appear less reliable, with some companies lagging several months in their reporting. I use the reports through 1985 to identify installations that took place during or before 1984.

show the locations of scanners in the 48 contiguous states as of December 1977 and December 1982.

For the period 1976–1984, the Longitudinal Business Database (LBD) maintained by the Census Bureau lists, with few exceptions, all business establishments with paid employees in the United States. (Exceptions include most government-owned or -operated establishments, establishments operated by religious organizations or schools, and agricultural establishments.) The LBD is described in detail in Jarmin and Miranda (2002), but I describe the most relevant features here. The LBD records the existence of an establishment and provides the establishment's industrial classification. Establishments are linked over time through a unique identification number, and are matched to the owning firm, which may be a single-establishment firm or a multi-establishment firm. In the retail context, a multi-establishment firm is usually a chain, although it can also be a firm operating one retail outlet and one or more non-retail outlets (manufacturing facilities, warehouses, etc.). Since scanners were initially installed exclusively in grocery stores and supermarkets, I use only establishments coded with SIC 541: grocery stores, food stores, and supermarkets.

My analysis critically depends on correctly matching scanner installation with stores in the LDB. The name(s) and locations (city and state) of all establishments in the LBD are obtained from the Business Register (BR), which relies on a mix of administrative records (e.g., tax filings) and Census collections (surveys and censuses). I match these to the FMI data, which include store name, city and state, with the LBD/BR. To maximize the match rate, I spent considerable effort standardizing names of stores and cities and correcting errors in spelling and geography in both datasets. Despite these efforts, only 3,683, or about 35%, of new scanner installations through 1984 in the FMI data can be matched with certainty to stores in the LBD.

Non-matching stores fall into three categories. Despite my efforts to standardize city names (e.g., "Saint" vs. "St."), a small number (a couple of hundred) of stores in the FMI data cannot be matched because the city in which they are listed does not match any known

city in the LBD. Of the remaining non-matches, approximately half do not match because the FMI publication lists only the store name, city, and state, and there are multiple possible matches in the LBD for that store. (Whenever possible, I use the store number to resolve such matches.) The other half do not match because the LBD does not include a store of that name in that city, due to a variety of data issues. First, store names vary due to differences in spelling, hyphenation, etc. I have made efforts to correct typos and standardize names, but inevitably, some differences remain. Mismatches also occur because stores change their names or are called by different names in the two files, for example because the legal name of the business, which appears in the LBD, differs from the store's name. Second, there are some mismatches due to having different cities listed in the two files, either because the cities are adjacent and the store is located near the boundary between them, or, in some cases, because the FMI file lists the nearest large metropolitan area and not the physical location of the store. Whenever the discrepancy can be resolved unambiguously, I reassign the city based on the BR information. Third, some stores with scanner installations are not coded by Census as food stores (SIC 541) and do not match for this reason. Fourth, since the LBD includes only establishments with paid employees, any non-employers installing scanners cannot be matched. (This explanation is unlikely to play an important role in the current application, because most installations appear to be in relatively large stores.) Finally, some scanner installations listed as new installations in the FMI data were actually upgrades of existing installations, implying that the FMI data overstate the extent of the technology's diffusion in the early years.

Figure 4 shows the number of installations per year in the matched sample. The match rate ranges from 27% (in 1978) to 42% (in 1982). Table 1 presents summary statistics for this sample. For each year, the table presents the number of observations for stores both with and without scanners, and compares the scanning stores to the non-scanning stores on three dimensions: annual revenue and payroll (deflated to 1982 real dollars using the all-items CPI) and productivity, measured as the ratio of the two. In both 1977 and 1982,

scanning stores were larger than non-scanning stores both in terms of their revenues and in terms of their employment. Kolmogorov-Smirnov tests reject the hypothesis that these observations are drawn from the same distribution, at least with respect to revenue, although it is impossible at this point to say whether this difference is due to selection or to the impact of the scanners. Despite this statistical difference, the distributions are largely overlapping and similar. Figure 5 shows the distributions of 1977 and 1982 log revenue by scanner status for this sample.

The LBD has information on establishment payroll and the number of paid employees, but does not include revenue or other output measures, so studying the productivity effects of scanners requires matching to a third data source: the Census of Retail Trade (CRT). The CRT is part of the Economic Census, which takes place quinquennially in years ending in "2" and "7." The micro files for the CRT are available electronically starting in 1977, so I am able to match the 1977 and 1982 CRT with the LBD/FMI data. Although the full 1972 CRT is not available electronically, the 1977 CRT provides the 1972 employment, revenue, and payroll for continuing establishments. The CRT provides employment as of the week of March 12, annual payroll, and annual revenue for each retail establishment. From these I calculate labor productivity as the ratio of revenue to payroll. To reduce the impact of outliers on coefficient estimates, I drop the top and bottom 1% of productivity values for each year.

I focus on labor productivity because the stated goal of most store managers was to reduce labor costs. Labor productivity is preferable to raw labor costs, however, because raw costs could have increased if sales increased. I prefer measuring productivity as the ratio of revenue to payroll over the alternative of revenue per worker for several reasons. First, both revenue and payroll are annual measures, whereas employment is given at a point in time. Second, revenue and payroll are measured in the same units, so payroll provides a natural deflator for revenue. Third, the employment figures do not distinguish between full-time and part-time workers, which introduces noise into this measure and may bias

coefficient estimates. If stores respond to scanner installations by, for example, increasing the number of full-time workers relative to part-time workers, we may see an effect on revenue per worker while revenue per hour is unaffected; using revenue per payroll dollar should eliminate problems of this sort. This is a particular concern since cashiers tend to work fewer hours than other grocery-store workers, so a reallocation of tasks across the store is likely to affect the distribution of hours. Finally, employment figures do not distinguish between high-skilled and low-skilled workers or managers and cashiers, and stores may change the mix of these as they install scanners, as well. I explore these issues in Section 5.1.

Even the ratio of revenue to payroll is not an ideal measure of productivity. An obvious problem is that wholesale costs are not netted out of revenue because the Census does not collect input costs from retailers at the store level.[4] Another problem is that revenue is sensitive to prices. The productivity gains I measure are therefore biased downward if stores lowered prices, or alternatively if wages increased, in response to scanner-induced increased worker productivity. In the absence of store-level price deflators there is no perfect way to solve this problem, but I am able to get at the magnitude of this problem in a specification check in which I test for a relationship between scanner installations at other stores in the city and a store's own productivity.[5]

A final problem is that much of the wage variation across establishments is known to reflect relative bargaining positions and other factors that do not reflect either hours of work or worker skill. For example, Abowd, Kramarz, Margolis, and Troske (2001) find substantial variation in wages among full-time workers even after controlling for skill. Using data from the Current Population Survey's Outgoing Rotation Groups for 1979–1982, I find that hours worked and highest grade attained explain 70–75% of usual weekly earnings by

[4]These data are collected in an annual survey at the firm level, but only published sector-level averages are available for this time period.

[5]All these problems are shared by the employment- (rather than payroll-) based productivity measure used by Foster, Haltiwanger, and Krizan (2002) and Doms, Jarmin, and Klimek (2004). See Foster, Haltiwanger, and Krizan (2002), Haskel and Sadun (2009), and Betancourt (2005) for further discussion.

grocery-store employees, depending on the year. Adding occupation (e.g., cashier, manager, meat cutter, etc.), state, and central-city controls increases the regression R^2 to 78–83%. Standard explanations for differential returns on skill focus on firm-level variables (typically unobservable). Since all my regression equations include store fixed effects, unobserved store- or firm-level variables should not bias the results in this paper as long as they do not change over the time period studied.

4 Scanners and Productivity

4.1 Difference-in-Difference Estimates

I estimate productivity regressions using a difference-in-difference specification:

$$\ln(\mathbf{productivity})_{it} = \alpha_i + \delta_t + \beta \mathbf{Scanner}_{it} + \varepsilon_{it} \tag{1}$$

where $\mathbf{productivity}_{it}$ is revenue per payroll dollar in store i in year t, α_i is a store fixed effect, δ_t is a time fixed effect, and $\mathbf{Scanner}_{it}$ is an indicator for the store having had a scanner for the full calendar year, or, if the scanner was installed partway through year t, the fraction of year t during which the store had a scanner. (I assume that a store that installed a scanner in January had the scanner for $\frac{11.5}{12}$ of the year, February installations were active for $\frac{10.5}{12}$ of the year, and so on.) Standard errors ε_{it} are clustered at the store level to allow for arbitrary autocorrelation in the error term.

I limit the analysis to the matched sample to address both measurement error and omitted-variable bias. Measurement error in the full sample of stores results from the fact that more than 6,000 stores listed in the FMI data as having installed scanners over the period of study cannot be matched to stores in the Business Register and Census data. Excluding stores in the Census that did not match with installations in the FMI data ensures that stores coded as not having scanners as of year t really do not have scanners. Otherwise,

measurement error in the scanner variable would cause attenuation bias of the coefficient β. Discarding all unmatched stores eliminates the measurement-error concerns.

Omitted-variable bias is present if the selection into scanning is not random, specifically, if it is either caused directly by store productivity or correlated with other factors that affect productivity. Because stores that adopted scanners in the early years were on average larger, more productive, and had higher productivity growth than non-adopting stores, including non-adopters in the regression would bias the estimate of β upwards. In Appendix A, I provide estimates using the full data set and show that the endogeneity bias dominates the measurement-error bias in the full sample.

Restricting my sample to matched stores imposes only weak conditions for interpreting the estimates causally. The estimates require only that, conditional on installing a scanner by December 1984, the exact timing of a store's installation is uncorrelated with the error term, so that stores that installed in 1983 and 1984 are valid controls for stores that installed scanners earlier. Endogeneity bias could still be present if stores that installed scanners in the late 1970s and early 1980s are systematically different from stores that installed scanners in 1983 and 1984, but these differences are likely to be much smaller than the differences between the "treatment" stores and stores that lacked scanners at the end of 1984. In the next section, I perform two specification tests to determine the extent to which these problems are in fact addressed with the limited sample.

My estimate of β is shown in the first column of Table 2. It shows a 4.5% increase in the productivity of stores that installed scanners by the end of 1982. Estimates are robust to several specification tests (not shown), including controlling for state-by-year fixed effects, to allow different trends across states, as well as for allowing year fixed effects to differ for stores in counties with 1970 population above 500,000. The result is also robust to adding firm fixed effects in addition to store fixed effects, to account for the small number of stores that change ownership over the sample period. The estimate continues to be statistically significant at the 1% level when clustered at the firm, rather than store, level. Estimates

using different samples, e.g., only supermarkets (the primary installers of scanners over this time period), are also very similar.

In Appendix B, I present alternative estimates using a spline method that exploits the non-linearities in expected productivity gains between 1977 and 1982 from installing a scanner in 1981, 1982, and 1983. This method, which uses significantly fewer observations (1,580 in total), generates an estimate of 4.6% productivity gain.

4.2 Specification Tests

In this section, I look for evidence of omitted-variable bias related to other productivity shocks. For example, productivity-enhancing management or ownership changes, or changes in store organization, could be concurrent with, or followed closely by, the adoption of scanning technology. One symptom of this sort of omitted-variable bias could be a positive "preemptive" estimate of scanner adoption on productivity. Specifically, stores that do not yet have scanners in year t but that install them shortly thereafter (in year $t+1$) may exhibit high productivity already in year t. To investigate the extent of this problem, I estimate

$$\ln(\mathbf{productivity})_{it} = \alpha_i + \delta_t + \beta_t \mathbf{Scanner}_{it} + \beta_{t+1} \Delta\mathbf{Scanner}_{i,t+1} + \varepsilon_{it}, \qquad (2)$$

where $\Delta\mathbf{Scanner}_{i,t+1}$ is equal to one if store i installed scanners between January and December of year $(t+1)$, and the other variables are as defined earlier. Column (2) of Table 2 shows estimates of β_t and β_{t+1}. The estimate of β_t is very similar to the one from Equation (1). The estimate of β_{t+1} is small and not statistically significant, consistent with random timing of installations, or at least with the timing of installations being uncorrelated with other factors affecting store productivity. In contrast, Appendix Table A-1 shows that if I do not restrict the sample to stores that installed scanners by 1984, β_{t+1} is estimated to be both large and statistically significant. Using the restricted sample appears to solve this omitted-variable bias.

A second specification test concerns local market conditions. If scanners installations coincided with, or caused, falling wages or increasing grocery prices, then estimates of the effect of scanners on productivity would be biased upward. (Conversely, estimates would be attenuated if installations coincided with increasing wages or decreasing prices.) This is a particular concern if faster-rising prices in some cities induced supermarkets in those cities to install scanners sooner than supermarkets and grocery stores in cities with lower food-at-home inflation rates. To test whether this effect is biasing my results, I create a new variable, **CityScanner**, which equals 1 if at least one *other* grocery store in the city had a scanner (new or pre-existing) in year t:

$$\mathbf{CityScanner}_{it} \equiv \begin{cases} 0 & \text{if } \max_{j \in c(i)/i} \mathbf{Scanner}_{jt} = 0 \\ 1 & \text{if } \max_{j \in c(i)/i} \mathbf{Scanner}_{jt} > 0 \end{cases}$$

where the set $c(i)/i$ denotes stores in store i's city, $c(i)$, excluding store i. To determine whether any grocery store in city c had a scanner in year t, I first remove a small number of stores from the FMI dataset which are clearly not grocery stores (mainly drugstores and general-merchandise stores). I include all other stores in the city, whether or not I can identify them in the Business Register. I then estimate

$$\ln(\mathbf{productivity})_{it} = \alpha_i + \delta_t + \beta \mathbf{Scanner}_{it} + \gamma \mathbf{CityScanner}_{it} + \varepsilon_{it}. \tag{3}$$

Results are shown in the last column of Table 2. Again, the estimate of β is virtually unchanged, while the estimate of γ is very small and not statistically different from zero, consistent with an installation process that is neither driven by, nor correlated with, city-level changes in prices or wages. Restricting the sample to stores without current scanners ($\mathbf{Scanner}_{it} = 0$) to avoid confounding the effects of store- and city-installed scanners, not shown, produces even smaller estimates of γ.

Based on the results of these specification tests, this sample appears to be free of endogeneity bias.

5 Understanding the Productivity Effect

5.1 Decomposition

In this section I separate the impact on store productivity, measured as the revenue-to-payroll ratio, into its components, estimating separately the effect of scanners on store revenue and payroll. I re-estimate Equation (1), replacing the left-hand side variable by $\ln(\textbf{revenue})_{it}$ and $\ln(\textbf{payroll})_{it}$, respectively. These results are reported in the first two columns of Table 3.[6]

Selmeier (2008, p. 238) reports that a month after installing scanners stores often experienced a 10-12% increase in revenue, but I find no evidence of such an effect. The point estimate of the effect of scanners on revenue is small (and negative), and although the confidence interval is large, a 10% increase in sales can be rejected at the 1% level. The effect on payroll, however, is negative, statistically significant, and, at -5.3%, large enough to fully account for the productivity increase estimated in the previous section.

Since the average annual payroll of stores in the sample is approximately \$816,000 (in 1982 dollars), these results imply that, on average, barcode scanners increased productivity by reducing payroll by about \$36,500 per year, and holding revenue unchanged.

For completeness, the rest of the table reports, respectively, the effect on March 12 employment, revenue per March-12 employee, and payroll dollar per March-12 employee. For these three variables, I also estimate an alternative specification in which I replace the scanner variable with an indicator which equals 1 if the store installed a scanner no later than March of year t; results (not shown) are unchanged. The results show that stores did not adjust their employment in response to the scanner installations. Since revenue was

[6]The number of observations in this table varies with the left-hand side variable because I remove the top and bottom 1% of each variable separately before estimating the regressions.

unchanged, too, productivity figures based on revenue per March-12 employee show no effect. But stores did significantly reduce their payroll outlays, by reducing the average number of hours worked per employee and/or the average hourly wage.

CRT data do not allow me to distinguish between these two channels, but Bloom (1972, p. 223) notes that one advantage of the new technology was that "the automatic feature of the operation would make it possible for stores to hire less competent applicants," suggesting the hourly wage may have decreased. This hypothesis is consistent with wage trends in the Current Population Survey (CPS) Outgoing Rotation Groups (ORG) from 1979 to 1982. The CPS ORG includes both a detailed industry code (equivalent to a 3-digit SIC code) and a detailed occupation code. Over these four years, nearly 25,000 respondents reported working in the grocery-store industry (1970 industry code 628 or 1980 industry code 601, equivalent to SIC 541); of these, 29.7% provided an occupation code of cashier (1970 occupation code 310 or 1980 occupation code 276).[7] Between 1979 and 1982, the real hourly wage of cashiers in this sample, deflated using the all-items CPI, fell by 9.9%, while the real hourly wage of other occupations fell, on average, by approximately 8.5%, lending additional plausibility to the skill hypothesis.

Of course, other variables, not measured by the Census, may also have been affected by scanner installation: for example, the cost of goods sold, or the cost of intermediate inputs. As noted earlier, the CRT does not contain any information on wholesale costs. None of the sources I have consulted for this study indicate there was an increase in wholesale costs despite the higher cost of UPC printing by food manufacturers, perhaps because that cost was fixed and not variable.

[7]Other common occupation codes are for stock handlers (23%), managers (14%), sales clerks (8%), and meat cutters and butchers (5%).

5.2 The Value of a Barcode

The benefits from scanning depends critically on the number of products labeled with Universal Product Codes. The more manufacturers adopted the UPC, the more retailers benefited from scanners. As Bloom (1972) notes, "If the operator at the checkout must stop and manually enter the code for items which are not marked, the benefits described could be substantially reduced" (p. 220).

Ideally, I would have liked to directly estimate the increase in productivity from having one more product bearing a UPC symbol. In the absence of supplier-level information on UPC adoption I cannot hope to estimate this effect at the store level. Instead, in this section I estimate the effect of carrying more packaged (e.g., boxed, canned, jarred) goods, since these products generally got UPC codes sooner than frozen and fresh products. I then use these estimates to back out the impact of one additional barcoded product on a store's productivity.

To test whether supermarkets whose revenue depended more heavily on products likely to be barcoded early saw bigger productivity gains from the introduction of scanners, I use a subset of food stores that provided an accounting of their revenue sources on their 1977 Census form.[8] Stores were asked to report the share of their revenue from several grocery categories (meat, fish, and poultry; fresh produce; packaged frozen foods; dairy products; bakery items; and all other grocery products), as well as alcohol, tobacco products, health and beauty aids (including prescription and non-prescription drugs), and other, less common, categories. I limit the sample to stores whose 1977 reported revenues in the various grocery subcategories sum to within 1% of their total grocery revenue, and whose reported combined revenues from food and all non-food product lines sum to within 1% of their total revenue. While the sample is relatively small, it allows me to estimate the extent to which differences in product emphasis affected stores' ability to exploit the new technology.

[8] I do not have 1972 records on product line sales, so I omit 1972 data from these regressions.

I estimate

$$\ln(\mathbf{productivity})_{it} = \alpha_i + \delta_t + \gamma_t \times \mathbf{packaged}_i + \beta \mathbf{Scanner}_{it} + \beta_l \mathbf{Scanner}_{it} \times \mathbf{packaged}_i + \varepsilon_{it} \quad (4)$$

where **packaged** is the share of store i's 1977 revenue obtained from packaged grocery products, which I define as excluding fresh produce and meats. The interaction term $\gamma_t \times \mathbf{packaged}_i$ captures the possibly different time trends of productivity for stores that specialize in packaged goods and stores that specialize in other food products, such as fresh produce.

Results are presented in the second column of Table 4. While the main effect β is negative, the coefficient β_l, which captures the additional productivity gain from scanning due to a higher share of packaged goods, is large, above 0.21, and significant at the 1% level. To interpret the results, note that across stores in the sample, the 25th percentile of **packaged** is approximately 0.5, and the 75th percentile is 0.7; the mean is 0.6. The 0.21 point estimate implies that a store that received half of its revenue from packaged groceries in 1977 experienced a productivity increase, on average, of only 1.4% when it installed a scanner, compared with a 5.7% productivity increase for a store that got 70% of its revenue from packaged groceries.

Making inferences from these estimates to the productivity gain from each additional UPC requires some additional assumptions. First, I assume for this discussion that packaged goods are synonymous with barcoded products. In practice, the diffusion of UPCs took time, and not all packaged goods were barcoded during this period. Conversely, some non-packaged products may have been early adopters of UPC symbols. Second, I use a store's share of revenue obtained from packaged goods as a proxy for the store's share of items with barcodes (effectively assuming that the average quantity-weighted price of packaged and non-packaged goods is the same). Food Marketing Institute reports cited by Messinger and Narasimhan

(1995, Table 6, p. 200) gives the average number of stock-keeping units, or products, in a supermarket in 1975 as 9,000, a number that increases to 14,000 by 1980. I assume the average store in my sample had 10,000 products when it introduced a barcode scanner.

With these assumptions, the estimates above can now be translated into the impact of a single barcode. Adding one barcoded product to a store with 6,000 barcoded products, in 1977, would have increased a store's productivity gain from installing a scanner by 0.0035%. Given the estimate in Section 5.1 that a 4.5% productivity gain saves the average store $36,500 in payroll annually, a single additional barcoded product increases the savings due to a scanner by approximately $28 per year per store.

Since the gain from barcoded products accrued to the retailer, not the manufacturer, this ballpark figure puts some bounds on the magnitude of network externalities that were central in the implementation of scanning technology. This figure should be interpreted with some care, however, since, at best, it represents an average, per product, per store. The frequency of purchase and the number of scanning stores carrying each product both directly affect the value of an individual barcode.

6 Cost-Benefit Analysis

The cost of scanners varied over time and by manufacturer. Brown (1997, pp. 80-81) gives a cost of "between $50,000 and $250,000 for a supermarket with $60,000 a week in sales" in 1973. In 1978, a *Time* magazine article estimated "the purchase price [of] a sophisticated eight-lane check-out system" at "more than $110,000" (Time Magazine, 1978). Consistent with this estimate, Shaw (1977, pp. 58-60) quotes prices from IBM for a 10-lane scanning system at $133,100 in 1976, not including maintenance fees which averaged $7,000/year. Shaw (1977) also provides other vendors' prices, which were roughly similar. Finally, Bloom (1972, p. 224), citing an early McKinsey & Co. estimate, predicts that the equipment would cost "in excess of $150,000" for a store with $4 million in annual revenue, and approximately

$100,000 for a store with half that revenue. These numbers provide a benchmark against which to check the credibility of the empirical estimates of scanners' productivity benefits.[9] Extrapolating to the larger stores that actually installed the early scanners, the cost is likely to have exceeded $200,000 in 1977 dollars, or about $300,000 in constant 1982 dollars, even accounting for a 10% decrease in the real price over this time period.

Like other major innovations with network externalities and social returns to scale, the scanner platform required players in the market to overcome two important barriers: coordination and short-term costs. (For background, see Bresnahan and Greenstein, 1999, .) If the primary barrier were one of coordination costs, we might expect profits to rise quickly once the platform is adopted. Instead, the above calculation suggests that costs were high relative to benefits in the short run, consistent with profitability limiting the rate of diffusion, with a more limited role for coordination costs, at least once the key design choices were made).

The opportunity cost of setting up a scanner included more than the cost of hardware and software. If the installation required the attention of managers be diverted from other tasks, as many of the early accounts of scanner adoption suggest, the short-term productivity gains could have been significantly lower than the above estimates suggest. To estimate this effect I separately estimate the short-term (current-year) and medium-term productivity impact of having a scanner installed.

The degree to which first-year gains are reduced can be captured by estimating

$$\ln(\mathbf{productivity})_{it} = \alpha_i + \delta_t + \beta_t \mathbf{Scanner}_{it} + \beta_s \mathbf{1}\left(\Delta \mathbf{Scanner}_{it} > 0\right) + \varepsilon_{it}, \qquad (5)$$

[9] Very different numbers are provided by Zimmerman (1999, p. 39), who quotes a former NCR employee saying scanners listed for "$4,200 in 1974 and decreased to about $3,800 by 1980." These numbers are roughly consistent with an early, but undated, document published by scanner system manufacturer National Semiconductor claiming the cost differential between a mechanical register and an optical scanner in the $7,000-$11,000 range (National Semiconductor Corporation Systems Division, undated). I was unable to determine the source of this discrepancy.

where $\mathbf{1}(\cdot)$ is the indicator function. The coefficient β_s represents the non-pecuniary "setup cost": the one-time contemporaneous effect on productivity of a scanner installation on productivity.

Estimates are shown in the third column of Table 4. The productivity gain a store experiences in the year of scanner adoption is significantly smaller, both statistically and economically, than its gain in subsequent years. A scanner installed in January would have increased the variable $\mathbf{Scanner}_{it}$ by $\frac{11.5}{12}$ and increased the variable $\mathbf{1}(\Delta\mathbf{Scanner}_{it} > 0)$ by one, leading to a net productivity gain of 2.2%. Each year thereafter, the store's productivity relative to the rest of the industry is 4.6% higher than its baseline.

There are several possible reasons for non-pecuniary setup costs that may be reflected in productivity figures. Early accounts of scanner installations, for example by the former IBM employee Bill Selmeier, are full of details of meetings with store managers. These meetings could have taken the managers' attention off other issues, such as pricing and scheduling, reducing observed productivity. Other types of investments, such as reorganization, may have also interfered with early returns (Brynjolfsson and Hitt, 2003). In addition, scanners' early productivity gains may have been tempered by additional training and a learning-by-doing phase. Because relatively few stores installed scanners during a Census year, I was unable to separately estimate possible "learning-by-doing" effects and the one-time productivity loss due to installation.

The productivity gains due to scanners probably increased over time for at least three reasons. First, as discussed earlier, later vintages of scanners, released starting in 1979, were able to read smaller barcodes and had fewer reading errors. Second, the number of firms registered with the Uniform Code Council increased monthly, and the number of manufacturers with UPC registrations more than doubled between 1977 and 1982 (Dunlop and Rivkin, 1997, Figure 1, p. 4). Finally, manufacturers improved their training of store managers and cashiers over time. Bill Selmeier, who worked for IBM in the early years of scanners, recalls the feeling in 1974: "As much as everyone expected the systems to be

productive, somehow they weren't" (Selmeier, 2008, p. 159). IBM was able to provide better training to stores implementing its technology after a sending team to investigate the source of the problems.

Since my estimates provide the marginal effect of the average scanner installed over the period 1976–1982, and installations are heavily weighted towards the later part of this period, these estimates most likely do not apply to the earliest scanners. To test for a differential effect of later scanners, I estimate

$$\ln(\mathbf{productivity})_{it} = \alpha_i + \delta_t + \beta \mathbf{Scanner}_{it} + \beta_{82} \mathbf{Scanner}_{i,82} + \varepsilon_{it}, \tag{6}$$

Estimates are shown in the third column of Table 4. The estimate of β, which captures the productivity effect of a scanner installed by December 1977, is negative, but because it is imprecisely measured, the 1977 effect of a scanner is hard to determine from this regression. The differential effect of a scanner installed between 1977 and 1982, β_{82}, is also imprecisely estimated, but the point estimate is large (5.65%) and different from zero at the 5% level.

As already noted, for a store with $816,000 in annual payroll, an average productivity gain of 4.5% is the equivalent of $36,500 per year in saved payroll. (Figures are all in 1982 dollars.) The shelf-life of a scanner is somewhat hard to gauge from the available data because upgrades are not always properly classified, but at least 10% of the scanners installed by the end of 1983 were upgraded by the end of 1984, with the first upgrades listed as early as 1978. Given the high cost of scanners and their relatively short life, it seems doubtful that the early scanners were cost effective despite their large productivity impact (and in contrast to McKinsey & Co.'s forecasts).

This is not to say that scanners' long-run benefits did not outweigh their costs. Circumstantial evidence links scanners to the explosion in the number of products carried by supermarkets (Ellickson, 2007, reports that the number of distinct products per supermarket more than doubled between 1980 and 2004). Unfortunately, the data I use in this paper

cannot be used to test this hypothesis, since Census product codes are extremely broad and only packaged goods were reliably barcoded in the early years of the technology. If scanners did lead to increased store size, they may also have contributed to chains' increased market power in the supermarket industry. Basker, Klimek, and Van (forthcoming) provide evidence that over this period, chains with the biggest increase in the number of products they carried also opened more stores, suggesting a complementarity between product breadth and chain size. Larger chains have been able to exploit other sources of competitive advantage, including economies of scale.

The biggest barrier preventing immediate realization of scanners' benefits was the relatively small number of manufacturers printing UPC symbols on their products. This number increased from about 2,000 at the time of the first scanner's installation in 1974 to over 8,700 by the 1982 Census and nearly 13,000 by the end of 1984 (Dunlop and Rivkin, 1997). Although no data exists on the number of products bearing UPC symbols, this number probably increased at a slower rate, since early adopters probably had more distinct products than late adopters. A second, closely related, barrier was the inability of scanners to "read" barcodes that were small, wet, wrinkled, or damaged; as noted earlier, these technical barriers were overcome in stages, in 1979 and 1980. Later still was the problem of coding random-weight products, such as fresh produce, solved.

A second barrier was popular opposition among consumer and union groups to the removal of item-level prices, the individual stickers retailers placed on each package with the item's current price. Although retailers expected to realize significant saving by replacing so-called "item pricing" with a single label on the shelf ("shelf pricing"), this potential for saving was constrained in some states by law and in many others by popular sentiment.[10] In addition to the potential for significant reductions in labor costs, Nakamura (1999) and others also note that the eventual elimination of item prices increased the ease of price

[10]Levy, Bergen, Dutta, and Venable (1997) and Bergen, Levy, Ray, Rubin, and Zeliger (2008) make a compelling case that item pricing was costly to stores and, indirectly, to consumers.

changes, allowing retailers to experiment and learn about price elasticities.[11] Consumer arguments against removing the prices centered on the difficulty of checking one's receipt for errors, the difficulty of comparison-shopping, and the belief that supermarkets without item prices would raise prices more often.[12] Unions and labor groups were also very much opposed to item-pricing removal, fearing loss of jobs. Dunlop and Rivkin (1997) report that by March of 1976, three months before scanners were installed, "an understanding was reached between the labor and management of the [Joint Labor Management] Committee [of the Retail Food Industry]: Labor would not block the introduction of scanning systems and would not pursue item pricing legislation, and management would agree to leave prices on most retail products" (p. 26). This barrier eroded over time, but only very slowly. By 1989, only half of all stores had removed item prices (Walsh, 1993, p. 102).[13]

A final and important barrier to full utilization of the new technology was removed with the development and adoption of complementary technologies. The use of scanner-collected information on sales for inventory-management purposes, which provided a major benefit to both retailers and manufacturers, was neither immediate nor obvious to early adopters. According to Abernathy, Dunlop, Hammond, and Weil (1999), for example, inventory control was not among the initial benefits anticipated by scanner adopters. Scanner-enabled inventory keeping, in turn, allowed the development of electronic data interchange (EDI), including automated purchasing and payments, which helped retailers realize large cost savings (Abernathy, Dunlop, Hammond, and Weil, 1999). Only once retailers fully integrated these technologies into their operations could they realize the full productivity benefits of

[11] More generally, the cost of price changes could affect retailers' choice of pricing strategy, e.g., "every-day low pricing" (ELDP) vs. "promotional" ("Hi-Lo") pricing. See Ellickson and Misra (2008) for a discussion of pricing strategies.

[12] This concern may not have been unrealistic considering the high inflation rate during scanning's first years.

[13] In a previous version of this paper, I attempted to use variation in states' legal environment to estimate the impact of item-pricing regulation on the magnitude of the productivity effect. The fact that item pricing remained the rule even when the law allowed stores to remove these prices, however, limited the power of this exercise.

scanners.

7 Concluding Remarks

Much of the conversation about the so-called "ICT revolution" has focused on the 1990s and 2000s, often with only murky measures of "technology." This paper goes back in time to trace the effect of an early innovation — the introduction of barcode scanners by grocery stores in the late 1970s and early 1980s. Using a specific well-defined measure of technological innovation, and detailed store-level data on productivity, I find evidence of a direct causal relationship between technology adoption and productivity. To my knowledge, this is the first systematic evidence on the impact of any specific innovation on worker productivity in the retail industry.

I find that the early scanners increased store-level productivity by 4.5% on average, though the very earliest scanners probably had a much smaller impact, and setup costs substantially reduced first-year returns. Each additional barcoded product contributed, on average, $28 to the cost saving. Despite these impressive productivity gains, scanners were probably not a cost-effective investment in the early years.

The new platform required distributed investment, unequal economic burden, and uncertain profitability. It is a remarkable feat that retailers, wholesalers, food manufacturers, and scanner manufacturers were able to accomplish this transformation absent any government intervention and without massive transfers or cross-subsidies across players. Nevertheless, my calculations suggest that limited short-run profitability, more than the heterogeneity of costs and benefits and associated coordination costs, played a key role in limiting the adoption of this technology. Part of the explanation for the relative ease of coordination may be the relative homogeneity of the original Ad Hoc Committee. Another was, perhaps, a sense of urgency. If the industry failed to adopt a uniform code, individual supermarket chains were said to be likely to adopt different codes, the cost of which was likely to have been

far higher. These higher costs would have been incurred not by supermarkets, in terms of higher production costs of scanners and duplication of research and development, but also by manufacturers, which would have had to contend with placing many competing symbols on their packages.

Were scanners an example of capital deepening, or did they represent a profound shift in the way capital and labor are used — in other words, a change in total-factor productivity? Census data on the retail sector are too coarse to answer this question directly, but the path of progress after scanners' initial deployment suggests they represented more than capital deepening. Most store-level technological innovations since the 1970s have built upon the early scanners. Software and hardware upgrades now enable retailers to keep track of inventories and engage in electronic procurement and payments, reducing back-office labor costs. Inventory management may be further enhanced by the widespread implementation of radio-frequency identification (RFID), which will enable employees to "scan" items not in their direct line of vision. At the front end, the recent introduction of self-checkout — impossible before scanning — also appears to be reducing labor costs with hardly any additional capital investment.

A Measurement Error and Selection Bias

I restrict the sample used in this paper to stores that installed scanners by the end of 1984, and which I was able to match with observations in the Census Bureau's Business register, in order to address both omitted-variable and measurement-error biases. In this Appendix, I re-estimate Equations 1-3 using the full set of stores classified in SIC 541 to investigate the extent of selection bias and measurement error in the full sample of stores. The results are shown in Table A-1.

The omitted-variable concern is that stores adopted barcodes because of some unobserved factor, which either has itself a direct effect on store productivity, or is correlated with other unobserved factors affecting store productivity. Including store fixed effects in the regression controls for any time-invariant store-level differences in both observed and unobserved factors, but cannot account for time-varying effects. For example, stores that implement scanning may have grown, on average, relative to other grocery stores in this time period, or may have been located in neighborhoods that become relatively more prosperous. If these unobserved characteristics create differential productivity trends in stores that ultimately adopt scanners and stores that do not, a naïve regression would attribute the differential trend to the impact of scanners. In the main text, I argue that this problem is eliminated when I restrict the sample to stores that adopted scanners by December 1984. The specification tests support this argument. While omitted-variable bias could be either positive or negative, in this case it is almost certain to overstate the effect of a scanner.

The measurement-error concern is not present in the restricted sample by construction, since every store included in the sample is known to have installed a scanner. But because I am unable to match more than 6,000 stores listed in the FMI publication to stores in the Census, it is very likely that the full sample of stores includes some for which my scanner variable is set to zero actually had scanners for part of the sample period. This measurement error creates attenuation bias in the estimated coefficients.

In the full sample, the estimate of β from Equation (1), shown in the first column,

is nearly 0.15, or three times the magnitude of the estimate in the restricted sample, indicating large omitted-variable bias. Consistent with this conclusion, the estimate of β_{t+1} from Equation (2) is also positive, large, and significant. The statistical significance is not merely an artifact of the large sample. Scanner installation in year $t+1$ is associated with a 10% increase in store productivity in year t, an increase that amounts to two thirds of the estimated post-installation productivity increase. The positive and statistically significant estimate of γ from Equation (3) in the full sample, shown in column (3), most likely reflects measurement error.

B Spline Estimation

As a complement to the difference-in-difference estimates presented in the main part this paper, I also present a piecewise linear regression that exploits the timing of installations. Since most scanners were installed in the second part of my sample, I focus for this section on the years 1981–1983.

This analysis exploits the nonlinearity in the relationship between the timing of installation and the estimated impact on the store's measured productivity in 1982. Specifically, while the impact of scanners installed in 1981 should be fully captured in 1982 productivity, the impact of scanners installed in 1982 should be partially captured in 1982 productivity — more so, the earlier in the year the installation occurred — and the impact of scanners installed in 1983 should have had no effect on 1982 productivity. To test this hypothesis I estimate a spline regression. The left-hand side variable is a store's productivity growth from 1977 to 1982. I limit the sample to the 1,580 stores that installed scanners from January 1981 to December 1983 and for which I have both 1977 and 1982 productivity measures. This limited sample imposes even weaker identifying assumptions than the "installer" scanner used in the main part of the paper, namely that, conditional on a store installing a scanner over this 3-year period, the timing of installation is uncorrelated with other factors related

to the store's productivity growth from 1977 to 1982.

For each store, I define **ScanTime** to be the time elapsed, in months, between January 1, 1982 and the store's installation date. For example, if a store installed a scanner in March 1982, this variable equals 2.5; if the store installed a scanner in November 1981, this variable equals -1.5. Since my sample is limited to stores that installed scanners from January 1981 to December 1983, **ScanTime** takes on values from -11.5 to 23.5. To accommodate the nonlinearity discussed above, I create two knots in **ScanTime**, allowing changes in the constant term starting at $\mathbf{ScanTime} = 0$, separating stores that installed in 1981 from stores that installed in 1982, and at $\mathbf{ScanTime} = 12$, separating stores that installed in 1982 from stores that installed in 1983. I then estimate the piecewise linear relationship between the timing of scanner installation and productivity growth using a spline regression:

$$\begin{aligned}\Delta \ln(\mathbf{productivity})_{i,1982} = \alpha_0 + \beta_0 \mathbf{ScanTime}_i \cdot \mathbf{1}(\mathbf{ScanTime}_i < 0) \\ + \alpha_1 \mathbf{1}(\mathbf{ScanTime}_i \in [0, 12)) + \beta_1 \mathbf{ScanTime}_i \cdot \mathbf{1}(\mathbf{ScanTime}_i \in [0, 12)) \\ + \alpha_2 \mathbf{1}(\mathbf{ScanTime}_i \geq 12) + \beta_2 \mathbf{ScanTime}_i \cdot \mathbf{1}(\mathbf{ScanTime}_i \geq 12) + \varepsilon_i, \quad \text{(B-1)}\end{aligned}$$

where, as before,$\mathbf{1}(\cdot)$ is the indicator function, and $\Delta \ln(\mathbf{productivity})_{i,1982}$ is the difference between establishment i's 1982 productivity and its 1977 productivity.

The results are shown graphically as the dashed (green) line in Figure B-1. I have normalized the coefficients so that the point estimate for the effect of a scanner installed at $\mathbf{ScanTime} = 12$ (January 1, 1983) is zero. Dashed vertical lines indicate the knots. The results show a downward-sloping relationship between 1982 productivity gains and the time of installation for the full period from January 1981 to December 1982, followed by a slight upward-sloping relationship. Only the coefficient on the middle segment of the spline, however, statistically significant at the 10% level. That coefficient indicates that each additional month *without* a scanner reduces a store's productivity in 1982, relative to its

1977 level, by 0.29%, or 3.5% annualized.[14]

To increase the power and precision of these estimates, I also estimate an alternative specification in which I force the slopes of the first and last segments to be zero, and omit the level shifters between segments. The estimated coefficients from this regression are shown as the solid (orange) line in Figure B-1. Each month without a scanner reduces the store's 1982 productivity by 0.38%. Put differently, scanning for the full year increases the store's productivity growth by 4.6%, a figure remarkably close to the difference-in-difference estimate presented in Table 2.

[14]In unreported regressions, I also estimate variants of this regression in which I replace the left-hand side variable with the store's 1982 productivity (instead of productivity growth), both with or without controlling for 1977 productivity. Those results show larger effects, annualized to productivity gains of 7.6% and 5.2%, respectively. In model variants that allow discontinuities at the knots, the discontinuities are never jointly significant.

References

Abernathy, F. H., J. T. Dunlop, J. H. Hammond, and D. Weil (1999) *A Stitch in Time: Lean Retailing and the Transformation of Manufacturing — Lessons from the Apparel and Textile Industries*. Oxford University Press, New York and Oxford.

Abowd, J. M., F. Kramarz, D. N. Margolis, and K. R. Troske (2001) "The Relative Importance of Employer and Employee Effects on Compensation: A Comparison of France and the United States," *Journal of the Japanese and International Economies*, 15(4), Journal of the Japanese and International Economies.

Basker, E., S. Klimek, and P. H. Van (forthcoming) "Supersize It: The Growth of Retail Chains and the Rise of the "Big Box" Retail Format," *Journal of Economics and Management Strategy*.

Beck, J., M. Grajek, and C. Wey (forthcoming) "Estimating Level Effects in Diffusion of a New Technology: Barcode Scanning at the Checkout Counter," *Applied Economics*.

Bergen, M., D. Levy, S. Ray, P. H. Rubin, and B. Zeliger (2008) "When Little Things Mean a Lot: On the Inefficiency of Item-Pricing Laws," *Journal of Law and Economics*, 51(2), 209–250.

Betancourt, R. R. (2005) *The Economics of Retailing and Distribution*. Edward Elgar Publishing, Cheltenham, UK, and Northampton, MA.

Bloom, G. F. (1972) *Productivity in the Food Industry*. MIT Press, Cambridge, MA.

Bresnahan, T. F., and S. Greenstein (1999) "Technological Competition and the Structure of the Computer Industry," *Journal of Industrial Economics*, 47(1), 1–40.

Brown, S. A. (1997) *Revolution at the Checkout Counter: The Explosion of the Bar Code*. Harvard University Press, Cambridge, MA.

Brynjolfsson, E., and L. M. Hitt (2003) "Computing Productivity: Firm-Level Evidence," *Review of Economics and Statistics*, 85(4), 793–808.

Bucklin, R. E., and S. Gupta (1999) "Commercial Use of UPC Scanner Data: Industry and Academic Perspectives," *Marketing Science*, 18(3), 247–273.

Das, N., E. M. Falaris, and J. G. Mulligan (2009) "Vintage Effects and the Diffusion of Time-Saving Technological Innovations: The Adoption of Optical Scanners by U.S. Supermarkets," *B.E. Journal of Economic Analysis and Policy (Advances)*, 9(1), article 23.

Doms, M. E., R. S. Jarmin, and S. D. Klimek (2004) "Information Technology Investment and Firm Performance in U.S. Retail Trade," *Economics of Innovation and New Technology*, 13(7), 595–613.

Dunlop, J. T., and J. W. Rivkin (1997) "Introduction," in Stephen A. Brown, *Revolution at the Checkout Counter: The Explosion of the Bar Code*. Harvard University Press, Cambridge, MA.

Ellickson, P. B. (2007) "Does Sutton Apply to Supermarkets?," *RAND Journal of Economics*, 38(1), 43–59.

Ellickson, P. B., and S. Misra (2008) "Supermarket Pricing Strategies," *Marketing Science*, 27(5), 811–828.

Foster, L., J. Haltiwanger, and C. J. Krizan (2002) "The Link between Aggregate and Micro Productivity Growth: Evidence from Retail Trade," National Bureau of Economic Research Working Paper 9120.

——— (2006) "Market Selection, Reallocation and Restructuring in the U.S. Retail Trade Sector in the 1990s," *Review of Economics and Statistics*, 88(4), 748–758.

Greenwood, J., and M. Yorukoglu (1997) "1974," *Carnegie-Rochester Conference Series on Public Policy*, 46, 49–95.

Haberman, A. L. (2001) *Twenty-Five Years behind Bars: The Proceedings of the Twenthy-Fifth Anniversay of the U.P.C. at the Smithsonian Institution, September 30, 1999*. Harvard University Press, Cambridge, MA.

Harmon, C. K., and R. Adams (1984) *Reading between the Lines: An Introduction to Bar Code Technology*. North American Technology, Inc., Peterborough, NH.

Haskel, J., and R. Sadun (2009) "Entry, Exit and Labour Productivity in UK Retailing: Evidence from Micro Data," in *Producer Dynamics: New Evidence from Micro Data*, ed. by J. B. Jensen, T. Dunne, and M. J. Roberts. University of Chicago Press.

Holmes, T. (2001) "Bar Codes Lead to Frequent Deliveries and Superstores," *RAND Journal of Economics*, 32(4), 708–725.

Jarmin, R. S., and J. Miranda (2002) "The Longitudinal Business Database," unpublished paper, U.S. Census Bureau.

Jorgenson, D. W., M. S. Ho, and K. J. Stiroh (2008) "A Retrospective Look at the U.S. Productivity Growth Resurgence," 22(1), 3–24.

Levin, S. G., S. L. Levin, and J. B. Meisel (1985) "Intermarket Differences in the Early Diffusion of an Innovation," *Southern Economic Journal*, 51(3), 672–80.

——— (1987) "A Dynamic Analysis of the Adoption of a New Technology: The Case of Optical Scanners," *Review of Economics and Statistics*, 69(1), 12–17.

——— (1992) "Market Structure, Uncertainty, and Intrafirm Diffusion: The case of Optical Scanners in Grocery Stores," *Review of Economics and Statistics*, 74(2), 345–350.

Levy, D., M. Bergen, S. Dutta, and R. Venable (1997) "The Magnitude of Menu Costs: Direct Evidence from Large U.S. Supermarket Chains," *Quarterly Journal of Economics*, 12(3), 791–825.

Marrano, M. G., J. Haskel, and G. Wallis (2009) "What Happened to the Knowledge Economy? ICT, Intangible Investment, and Britain's Productivity Record Revisited," *Review of Income and Wealth*, 55(3), 686–716.

Mas, A., and E. Moretti (2009) "Peers at Work," *American Economic Review*, 99(1), 112–145.

Matsa, D. A. (forthcoming) "Competition and Product Quality in the Supermarket Industry," *Quarterly Journal of Economics*.

Messinger, P. R., and C. Narasimhan (1995) "Has Power Shifted in the Grocery Channel?," *Marketing Science*, 14(2), 189–223.

Nakamura, L. I. (1999) "The Measurement of Retail Output and the Retail Revolution," *Econometrica*, 32(2), 408–425.

National Semiconductor Corporation Systems Division (undated) *How to Buy a Supermarket Checkout System: The National Semiconductor Datachecker*. National Semiconductor Corporation.

Selmeier, B. (2008) *Spreading the Barcode: Personal Memories of Bill Selmeier*. Lulu.

Shaw, R. N. (1977) "Universal Product Code Scanning Systems: The Retail Experience 1974–1976," Ph.D. thesis, Cornell University, Department of Agricultural Economics.

Time Magazine (1978) "The Computer Society: Checking Out Tomorrow," *Time Magazine*, February 20.

Walsh, J. P. (1993) *Supermarkets Transformed: Understanding Organizational and Technological Innovations*. Rutgers University Press, New Brunswick, NJ.

Zimmerman, G. E. (1999) "The Effects of Network Externalities on the Adoption of the Universal Product Code: An Empirical Investigation of Grocery Retailers and Manufacturers, 1970–1980," Harvard University undergraduate honors thesis.

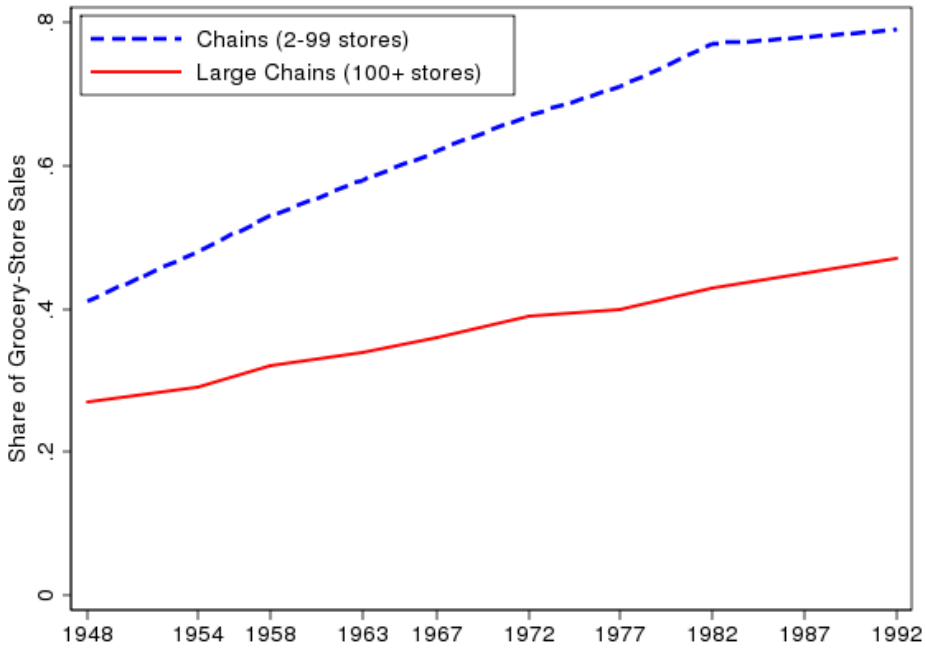


Figure 1. Chains' and Large Chains' Share of Grocery Revenues, 1948–1992
Source: Published data from Census of Business and Census of Retail Trade

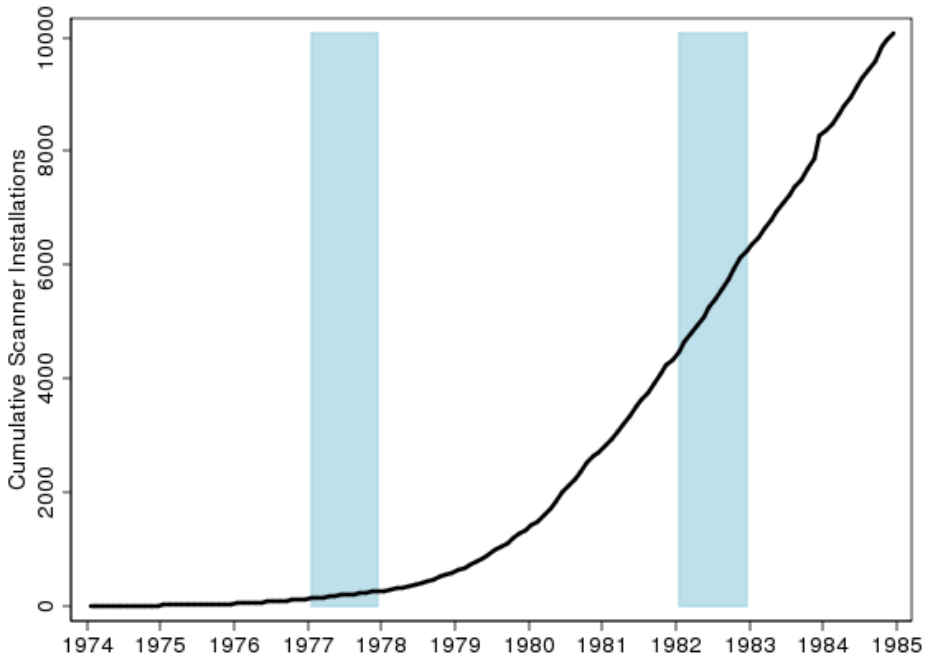
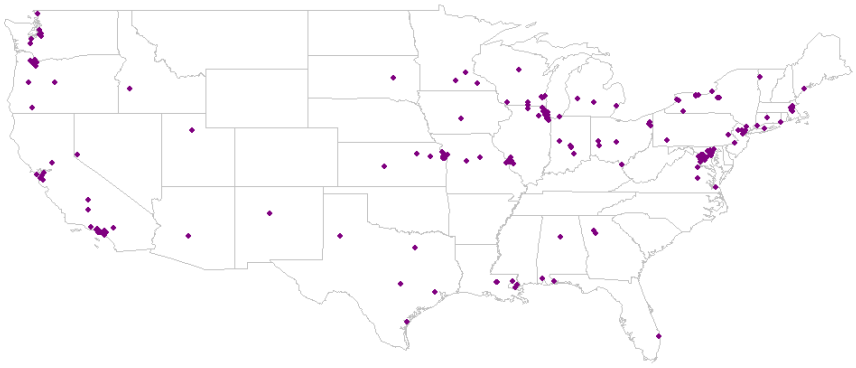
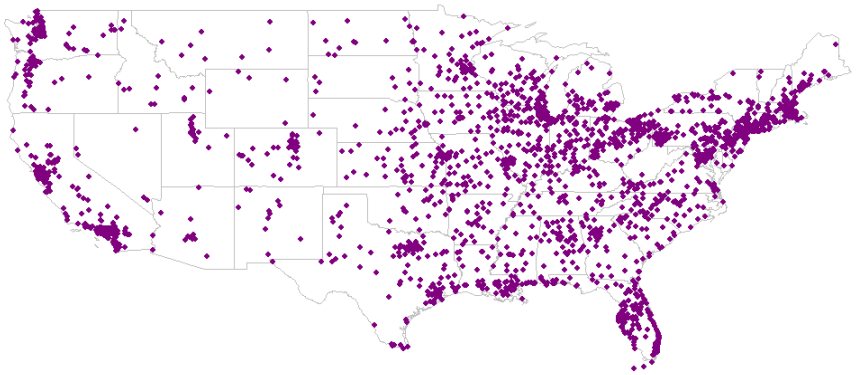


Figure 2. U.S. Scanning Stores, 1974–1984
Shaded periods represent Census years
Source: Author's calculations from Food Marketing Institute data



(a) 1977



(b) 1982

Figure 3. Scanner Locations, Year End 1977 and 1982

Alaska and Hawaii installations not shown
Source: Author's calculations from Food Marketing Institute data

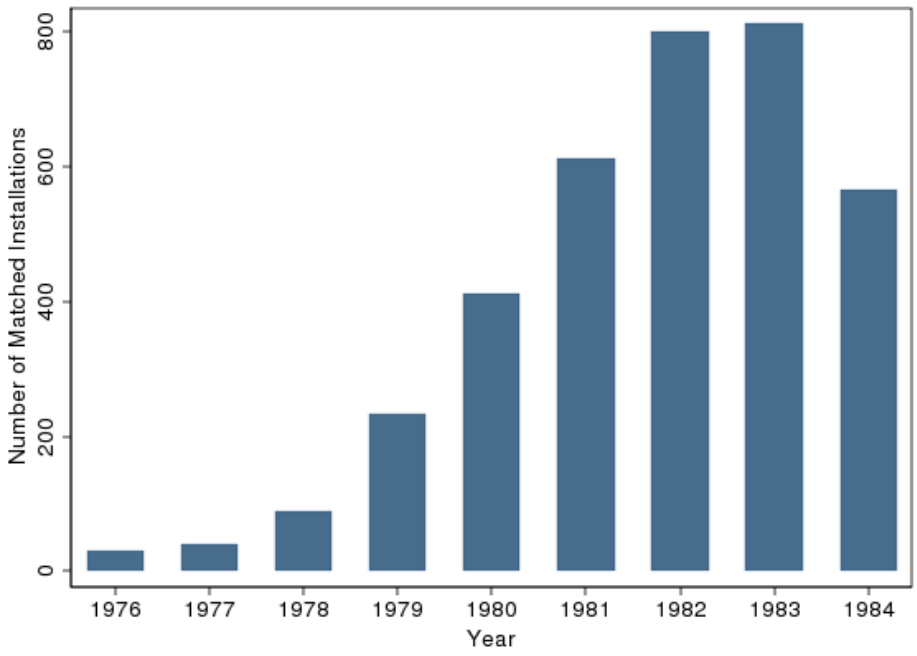


Figure 4. Number of Scanner Installations Matched to Business Register, by Year

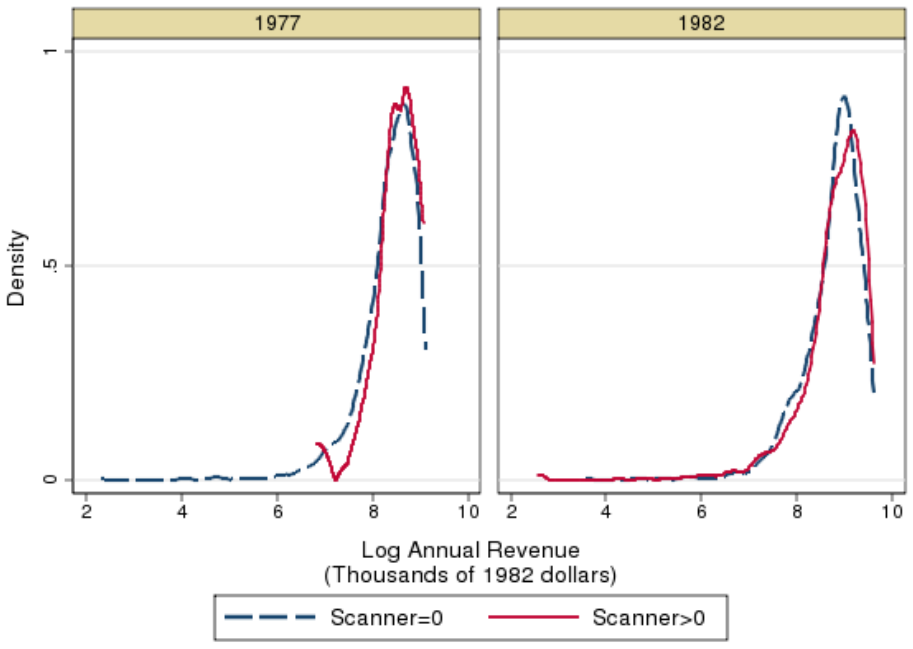


Figure 5. Log Annual Revenue by Current Scanner Status, 1977 and 1982
Sample of Stores that Installed Scanners by December 1984

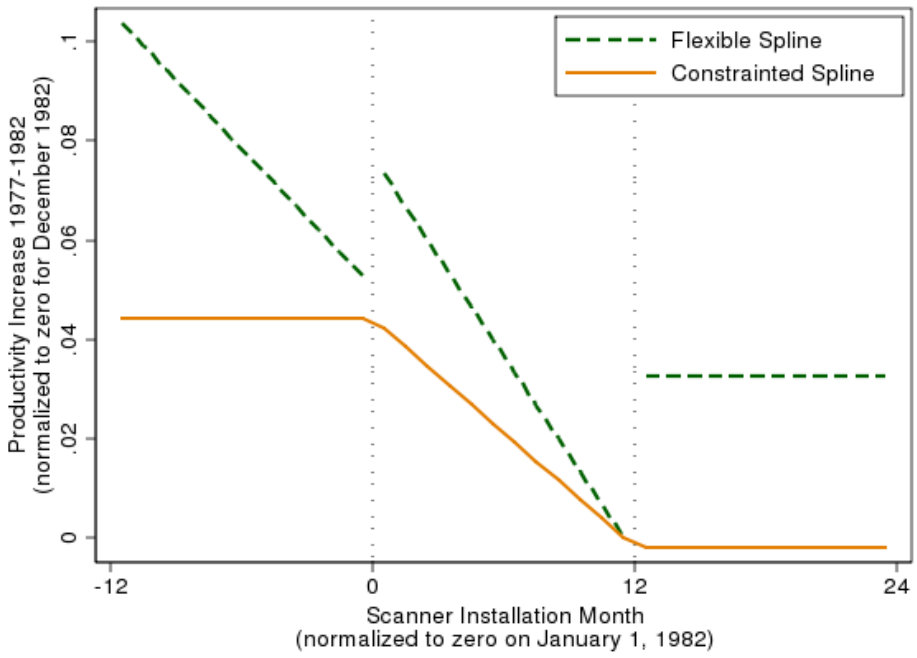


Figure B-1. Spline Estimates of Scanners' Effect on Labor Productivity

Table 1. Summary Statistics

		Stores	Average Revenue	Average Payroll	Average Productivity[a]
1972		1,698	8,004	711	12.0
1977:	**Scanner** $= 0$	2,624	8,958	832	11.4
	Scanner > 0	70	11,526	1,121	11.3
Kolmogorov-Smirnov[b]			0.047	0.015	0.011
1982:	**Scanner** $= 0$	1,355	8,208	810	10.7
	Scanner > 0	2,116	8,999	874	11.1
Kolmogorov-Smirnov[b]			0.000	0.114	0.000

Note: Sample of stores that installed scanners by December 1984
Revenue and payroll are in thousands of 1982 dollars
[a] Ratio of revenue to payroll; see text for details
[b] p-value from Kolmogorov-Smirnov test for equality of the distributions for stores with and without scanners each year

Table 2. Scanners' Effect on Labor Productivity: Difference-in-Difference Estimates

	(1)	(2)	(3)
Scanner	0.0451***	0.0495***	0.0454***
	(0.0084)	(0.0105)	(0.0084)
$\Delta\text{Scanner}_{t+1}$		0.0076	
		(0.0099)	
CityScanner			−0.0068
			(0.0073)
Year FE	✓	✓	✓
Store FE	✓	✓	✓
Observations	7,865	7,865	7,865

LHS variable is log productivity. Unbalanced panels, 1972–1982.
Robust standard errors in parentheses (clustered by store)
*** significant at 1%; ** significant at 5%; * significant at 10%

Table 3. Scanners' Effect on Alternative Outcomes: Difference-in-Difference Estimates

	Annual Revenue	Annual Payroll	March 12 Employment	Revenue per March Employee	Payroll per March Employee
Scanner	−0.0097	−0.0535**	0.0299	0.0022	−0.0546***
	(0.0253)	(0.0259)	(0.0256)	(0.0173)	(0.0210)
Year FE	✓	✓	✓	✓	✓
Store FE	✓	✓	✓	✓	✓
Observations	7,050	7,050	6,960	7,552	6,556

LHS variables in column headers (in logs). Unbalanced panels, 1972–1982.
Robust standard errors in parentheses (clustered by store)
*** significant at 1%; ** significant at 5%; * significant at 10%

Table 4. Scanners' Effect on Labor Productivity: Interactions

	Revenue from Packaged Goods[b]	New Installation[a]	1977 vs. 1982
Scanner	−0.0914**	0.0460***	−0.0072
	(0.0436)	(0.0085)	(0.0241)
Scanner × packaged[c]	0.2115***		
	(0.0718)		
$\mathbf{1}(\Delta\text{Scanner}_t > 0)$		−0.0234***	
		(0.0091)	
Scanner_{1982}			0.0546**
			(0.0242)
Year FE	✓	✓	✓
Year FE × packaged	✓		
Store FE	✓	✓	✓
Observations	4,913	7,863	7,863

LHS variable is log productivity.
Robust standard errors in parentheses (clustered by store)
*** significant at 1%; ** significant at 5%; * significant at 10%
[a] Unbalanced panel, 1972–1982; [b] Unbalanced panel, 1977–1982
[c] 1977 revenue share from packaged goods; mean value is approximately 0.6

Table A-1. Scanners' Effect on Labor Productivity: All Grocery Stores

	(1)	(2)	(3)
Scanner	0.1484***	0.1540***	0.1428***
	(0.0068)	(0.0069)	(0.0068)
ΔScanner$_{t+1}$		0.1000***	
		(0.0074)	
CityScanner			0.0330***
			(0.0038)
Year FE	✓	✓	✓
Store FE	✓	✓	✓
Observations	247,242	247,242	247,242

LHS variable is log productivity. Unbalanced panels, 1972–1982.
Robust standard errors in parentheses (clustered by store)
*** significant at 1%; ** significant at 5%; * significant at 10%