Upstream, Downstream:
Diffusion and Impact of the Universal Product Code*

Emek Basker  Tim Simcoe
U.S. Census Bureau  Boston University

June 2017

Abstract

The Universal Product Code (UPC) is widely viewed as a catalyst for the restructuring of retail supply chains during the 1980s and 1990s. Although casual observation reveals that scanners and bar codes are now ubiquitous, there remains very little quantitative evidence of their effects. This paper matches archival data from the Uniform Code Council to establishments in the Longitudinal Business Database and Economic Census to study the diffusion and impacts of the UPC. We find evidence of network effects in the diffusion process. Matched sample difference-in-difference estimates show that employment and sales increase following UPC adoption by manufacturers or wholesalers. Firms also apply for more trademarks after adopting UPC, consistent with the hypothesis that bar codes helped stimulate variety-enhancing product innovation.

JEL Codes: O33, L11, L60, L81

Keywords: Universal Product Code, Barcode, Supply Chain, Network Effects, Trademarks

*Comments welcome. Author contact: emek.m.basker@census.gov and tsimcoe@bu.edu. Any opinions and conclusions expressed herein are those of the authors 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. We thank Markus Mobius and David Weil for help obtaining the UPC data, Nathan Goldschlag and Nikolas Zolas for help with the trademark data and for their trademark-firm bridge file, and Randy Becker, David Brown, James Conley, Emin Dinlersoz, Lucia Foster, Nathan Goldschlag, Alex Krasnikov, Mark Kutzbach, Paul Messinger, Guy Michaels, Martha Stinson, Mary Sullivan, Kirk White, and seminar participants at the U.S. Census Bureau, MINES-ParisTech, the 2017 AEA (Chicago), and the 2017 IIOC (Boston) for helpful comments and conversations.
1 Introduction

The Universal Product Code (UPC) is widely touted as a major success of voluntary standardization. It was conceived in 1969 as a “standard human- [and machine-] readable code, to be used at all levels in the distribution channel” (Wilson, Jr., 2001, p. 2). The UPC has been credited with increasing product selection in stores (Holmes, 2001; Mann, 2001), shifting the balance of power along the supply chain from manufacturers to retailers (Messinger and Narasimhan, 1995), and stimulating labor productivity growth by enabling the growth of large retail chains (Sieling, Friedman, and Dumas, 2001; Foster, Haltiwanger, and Krizan, 2002). Tim Harford even named the bar code one of “50 Things That Shaped the Modern Economy.”

Although casual observation reveals that scanners and barcodes are ubiquitous in modern retail supply chains, there remains very little quantitative evidence of their effects. We link archival data on UPC registrations to firm-level data on employment, survival and trademark applications to provide new evidence on the diffusion and impacts of the UPC.

Previous accounts of UPC diffusion have emphasized that bar codes originated within the grocery industry before spreading to general merchandising and other retail supply chains (Dunlop and Rivkin, 1997). We examine the role of network effects within this diffusion process. Two-sided network effects imply that the return to adoption is higher for upstream firms supplying a “UPC-ready” downstream sector, and vice versa. Downstream firms become UPC-ready by installing scanners and by developing electronic data interchange (EDI) capabilities with their suppliers, including electronic payments (Abernathy, Dunlop, Hammond, and Weil, 1999; Basker, 2012). Although we do not have data on scanner installation, an indirect test of network effects is whether firms are more likely to register for a UPC.

---

1 For his explanation of this choice, see [http://www.bbc.co.uk/programmes/p04k0066](http://www.bbc.co.uk/programmes/p04k0066).

2 The UPC registration data come from the Uniform Code Council, and identify the date when a firm is assigned a unique identifier that comprises the first five digits of its code. This typically signals that the firm intends to add UPCs to its product labels. The outcomes data come from the U.S. Census (Longitudinal Business Database and Economic Census) and the U.S. Patent and Trademark Office.
when their rivals also register; in other words, whether UPC registration exhibits strategic complementarities within industry. We find strong evidence of strategic complementarities in UPC registration.

After examining UPC diffusion, we study the impact of manufacturer and wholesaler UPC adoption on three outcomes: employment, firm survival, and trademark registrations. Difference-in-difference regressions on a matched sample of UPC adopters and non-adopters show an immediate and significant increase in employment in the year after UPC registration. We discuss several possible mechanisms for this result, including that firms select into UPC registration due to anticipated demand shocks; that UPC registration proxies for adoption of a broader set of complementary technologies, such as EDI and inventory control systems, which can increase demand or lower costs (Hwang and Weil, 1998; Holmes, 2001); and that UPC registration may enable growth through business diversion if retailers prefer to work with upstream vendors that have adopted barcodes. The timing of the increase in employment is consistent with a combination of selection on positive demand shocks and business stealing by manufacturers and wholesalers that can integrate into large retailers’ supply chains more readily after registering for a UPC. We also show that the employment gain from UPC registration is increasing in the share of the firm’s competitors who have also registered, consistent with the presence of network effects.

A survival analysis shows that UPC adopters have a lower post-registration hazard of exit from the Longitudinal Business Database (LBD) than a sample of matched control firms with similar age and pre-adoption size and growth rates. Finally, we exploit a new link between the LBD and the USPTO Trademark Case Files Dataset (described in Dinlersoz, Goldschlag, Myers, and Zolas, 2017) to study the link between UPC registration and trademark applications, which we interpret as a proxy for variety-enhancing product innovation. Time-series data shows that within the grocery sector, growth in trademark applications, new product introductions, and the number of Stock Keeping Units (SKUs) stocked in a
typical supermarket all increased sharply in the early 1980s. Difference-in-difference regressions also show a substantial within-firm increase in the rate of trademark applications after a manufacturer or wholesaler registered for UPC. These results suggest that the supply-chain coordination enabled by the UPC system played an important role in stimulating increased innovation and new product variety.

This research contributes to a small empirical literature on bar codes and scanning. Within that literature, Basker (2012) is the most natural complement to this study. She examines the other half of the UPC-scanner platform, and finds that grocery stores experience a 4.5 percent increase in labor productivity after installing scanners. Basker (2015) finds that scanners also led to lower grocery prices. Dunlop and Rivkin (1997) examine the diffusion of the UPC using the same registration data that we exploit. By merging the UPC registration file to the LBD and Economic Census, we are able to examine the registration rates and shares (as opposed to a simple count). We present new stylized facts including that registration rates were strongly correlated with firm size, were higher on average in manufacturing than in wholesale (and in wholesale than in retail), and varied considerably by industry within the manufacturing sector.

This paper also contributes a literature on technology diffusion with network effects that is well summarized by Farrell and Klemperer (2007). One of the key points in that literature is that network effects (i.e., positive externalities amongst the users of different components in a system) can lead to excess inertia in technology adoption, as each side waits for the other to move first. One way to resolve this coordination problem is for interested parties to seek consensus on the technology, as well as the terms and timing of its adoption, within a Standard Setting Organization (Simcoe, 2012). This paper examines the diffusion of a standard after it is established, and in that respect resembles other studies that measure

---

3 SKUs are alphanumeric codes that track individual product data at a very granular level and used within a retail organization. UPCs, which can be used as SKUs, are standardized to allow for inter-firm communication and coordination.
network effects in two-sided platform adoption, such as Gandal, Kende, and Rob (2000) for Compact Discs and CD players, or Gowrisankaran, Rysman, and Park (2010) for Digital Video Discs and DVD players. The historical evidence suggests that for bar codes, the “chicken and egg” problem was solved by having manufacturers move first (at least within the grocery industry, which was first to adopt the system).

Finally, we contribute to a literature on the economic impacts of standards adoption. The empirical literature on this topic is relatively small, at least partly because it is hard to find data on standards diffusion. Simcoe and Rysman (2008) circumvent the data problem by examining patents that may be incorporated into a standard, and they find that these patents’ citation rates increase by roughly 20 percent following standardization. Bernhofen, El-Sahli, and Kneller (2016) examine the impact of the shipping container — a standard that is often mentioned in conjunction with bar codes — and find large increases in bilateral trade between countries that have each installed one or more container-ready ports. Gross (2017) examines a single large adoption event: the conversion of more 13,000 miles of U.S. railroad track to a standard gauge over one weekend in 1886. He finds that this led to a sizable redistribution of traffic from steamship to railroad on affected routes.

The rest of the paper is organized as follows. In Section 2 we provide general background on the Universal Product Code and its diffusion. We describe the data sources and our methods for combining them in Section 3. Section 4 presents some new facts about the diffusion process, including evidence consistent with the presence of network effects. In Section 5 we estimate the effect of UPC registration on upstream firms’ employment, survival and trademarking. Section 6 concludes.

2 The Universal Product Code

The Universal Product Code, originally the Uniform Grocery Product Code (UGPC), is a system of assigning a unique number to every product. It was initiated, designed, and
implemented by food-industry participants — manufacturers, wholesalers, and retailers — with no government oversight. Unlike the previous major retail innovation — mechanical cash registers, introduced in the 1880s — bar codes required standardization across the supply chain. The developers of the UPC expected that most benefits would accrue to retailers, but significant costs would be borne by suppliers (Brown, 1997, p. 114).

As designed by the Ad Hoc Committee on a Uniform Grocery Product Identification Code in the early 1970s, the barcode consisted of two five-digit numbers. The first five-digit number, a member prefix, was assigned by the Uniform Code Council (UCC) to paying member firms. Prefixes were purchased on a one-time basis at sliding-scale rates ranging from a couple of hundred dollars to over $10,000, depending on the revenues of the firm (Brown, 1997, pp. 119, 151). The second part of the code was assigned by the firm and could vary by product type, size, color, flavor, and even production date. Computer code associated each prefix with a manufacturer, and each suffix with a product and a price.\(^4\)

Registering for a UPC is necessary but not sufficient to placing barcodes on products. It is the latter innovation that enables scanning by retail outlets. Printing the barcode symbol required manufacturers to redesign their product labels to make room for the symbol, and in some cases to invest in printing technologies that allowed for sufficiently precise bars and minimized smearing. Importantly, our data allow us to determine when a company registered for one (or more) UPC prefixes, but not whether, when, or at what intensity it incorporated barcodes into its product labels.

Printing UPC symbols on packaging only benefits downstream firms to the extent that they utilize scanners. The adoption of scanners proceeded much more slowly than the UPC registration process. Brown (1997, p. 115) reports that by mid-1975, “50 percent (by volume)

---

\(^4\)Manufacturers were not limited to a single registration, and many large firms registered for multiple prefixes. Because a single prefix could accommodate up to 10,000 unique product codes, it is hard to imagine why more than a handful of firms would need multiple prefixes. In many cases, however, firms registered different prefixes to different divisions with the company (e.g., food and cosmetics divisions, or the Southern and Western divisions). Thus, the registration cost does not appear to have been prohibitively high.
of the items in a supermarket were source-marked with U.P.C. symbols, and thirty stores were actually scanning at the checkout counter.” Manufacturers expressed some dismay over the imbalance in adoption rates. For example, the July 1976 edition of *UPC Newsletter* contained an editorial entitled “Time to Get Involved” noting that there had been 4,412 manufacturer UPC registrations, compared to only 78 retail scanner installations (Uniform Product Code Council, 1976). One reason for imbalance is that a single UPC registration — sufficient for a firm with 10,000 individual SKUs — was much cheaper than installing scanners at the checkout.\(^5\) Food and beverage producers may also have been induced to adopt barcodes due to a Food and Drug Administration rule, adopted in 1973, that required nutritional information to be added to food labels, thus requiring a label redesign.\(^6\)

Although point-of-sale scanners initially diffused slowly at grocery stores, Figure 1 in Basker (2012) shows acceleration around 1981, so that by 1984 more than 10,000 supermarkets had installed a scanner.\(^7\) Around that time, the major general-merchandise retailers started using scanners in the back end of stores for inventory management, often in conjunction with early EDI implementations. For example, Kmart reported that in 1981 it implemented back-end scanners whereby “store employees use a wand to scan hardline merchandise on the sales floor and in the stockroom, assuring accurate replacement of goods” (Kmart, 1982, p. 9). Between 1982 and 1986, each of Wal-Mart’s annual reports makes some reference to investments in UPC-based point-of-scale scanning systems.

\(^{5}\)Basker (2012) estimates the cost of an early scanning system for a multi-lane supermarket at $300,000 in 1982 dollars.

\(^{6}\)The incremental cost of adding a barcode to a label during its redesign may have been relatively low (Brown, 1997, p. 62).

\(^{7}\)A series of papers by Levin, Levin, and Meisel (1985, 1987, 1992) and Das, Falaris, and Mulligan (2009) document the dynamics of scanner diffusion across grocery chains and metro areas in the U.S. A consistent finding in these papers is that there are positive spillover effects in scanner adoption within metro areas: supermarkets in areas with higher prior adoption are more likely to adopt scanners in later years.
3 Data

To study the diffusion and impact of the UPC, we constructed a firm-level panel dataset containing information on UPC registrations, employment and trademarking. These data contain information on 779,300 manufacturing firms and 866,500 wholesalers over the period 1975 to 1992, comprising 10.7 million firm-year observations. Table 1 reports summary statistics for each of the variables described below.\(^8\)

We use two source files to identify UPC registrations: a July 1974 membership list in the Uniform Grocery Product Code Council (Distribution Codes, Inc., 1974), and updated membership files used by John Dunlop in several papers (including Dunlop, 2001; Abernathy, Dunlop, Hammond, and Weil, 1995) and by Mobius and Schoenle (2006).

There are close to 100,000 registrations through 1992 in the Dunlop file. The left panel in Figure 1 shows the total number of registrations per year. The file also contains a “fee class” variable that indicates the price paid by the registering firm (Zimmerman, 1999, Appendix E). The fee class variable takes seven values, based on annual revenue (in millions of dollars).\(^9\) The right panel in Figure 1 shows the number of registrations by fee class; the vast majority are small firms with less than $2 million in sales.\(^10\)

We use name and address data in the UPC registration files to match registrants to business establishments in either the Economic Census (1977, 1982, 1987, and 1992) or the Longitudinal Business Database (1975 to 1992). Details of the matching procedure are described in the Data Appendix. Ultimately, we successfully match between 40 and 50 percent of UPC registrations to the Business Register. Figure 2 shows the match rate, by

---

\(^8\)Observation and firm counts are rounded to the nearest hundred to comply with Census rules on disclosure avoidance.

\(^9\)For example, firms with annual revenue up to $2 million were in fee class zero, and paid no more than $300 for their UPC registration; firms with annual sales above $500 million were in fee class 500, and paid $10,000 or more for a UPC registration. Zimmerman’s price list may be an updated version of an earlier list; Brown (1997, p. 119) pegs the initial fee floor at $200 and notes it increased to $300 later on.

\(^10\)The figure does not show the correlation between size classes and registration dates, but the share of size-class zero among the registrations increases sharply from 26 percent in 1972 to 86 percent 1978, after which it stabilizes between 85 and 90 percent.
year and by firm size. Not surprisingly, our match rate is disproportionately high for large firms, and therefore higher in the early years, when new registrants tended to be larger.

There are two main reasons for the match rate being bounded away from 100 percent. First, we only include firms with one or more establishments in three sectors: manufacturing, wholesale and retail. UPC registrations by firms in service industries, construction, transportation and warehousing, agriculture, and other sectors are omitted. Second, we omit non-employer businesses from our candidate matches. In addition, despite the large number of possible names and addresses we have in both files, there are inevitable mismatches that do not get resolved.

To create a panel data set, establishments are linked over time using the Longitudinal Business Database.\textsuperscript{11} However, because a UPC registration can be used by multiple establishments within a firm, we aggregate to the firm level using Census records that identify establishments with common ownership in any given year.\textsuperscript{12}

Table 1 shows that 3.8 percent of the observations in the manufacturers sample and 2.2 percent of the observations in the wholesalers sample belong to firms that registered for a UPC. These figures understate the diffusion of UPC, because of the highly skewed firm-size distribution, and the fact that larger firms were more likely to adopt, as we illustrate below. On average, the manufacturers in our panel are larger than the wholesalers, with mean employment of 72.8 and 14.9 persons respectively. There is considerable churn in both samples, with roughly 9.7 percent of firms exiting the data set in any given year.

To estimate network effects, we calculate UPC adoption by rivals in the same four-digit SIC code as a focal firm.\textsuperscript{13} Because firms may have one or more establishments (plants,

\textsuperscript{11}The LBD is described in detail in Jarmin and Miranda (2002). A key feature of the LBD is that it includes only business establishments with employees; therefore if an establishment continues to operate but without any employees, it appears as an exit in the LBD.

\textsuperscript{12}For additional details on the aggregation procedure, see the Data Appendix.

\textsuperscript{13}Four-digit manufacturing industries are generally more disaggregated than four-digit wholesaler industries. Examples of four-digit SIC codes in manufacturing are canned fruit and vegetables (2033), apparel belts (2387), and wood TV and radio cabinets (2517). Examples in wholesale are general-line groceries (5141), apparel piece goods and notions (5131), and furniture (5021).
warehouses, etc.), each with a different SIC, we construct our measure of rivals’ UPC adoption in two stages. In the first stage, we calculate the weighted average of UPC adoption by competitors for each establishment as:

$$\text{UPC}_{j(e)t} = \frac{\sum_{j \in \{s(e) \setminus i(e)\}} \text{Employment}_{jt} \text{UPC}_{jt}}{\sum_{k \in \{s(e) \setminus i(e)\}} \text{Employment}_{kt}}$$

where $\{s(e) \setminus i(e)\}$ is the set of all establishments with the same four-digit SIC as establishment $e$, but not owned by the firm $i$ that owns establishment $e$.\textsuperscript{14} This produces an employment-weighted average adoption rate of all rivals in the same four-digit SIC code as establishment $e$. In the second step, we take a weighted average of rival adoption rates across all of the establishments in firm $i$:

$$\overline{\text{UPC}}_{j(i)t} = \frac{\sum_{e \in i(e)} \text{Employment}_{et} \text{UPC}_{j(e)t}}{\text{Employment}_{it}}$$

Table 1 shows that employment-weight rival UPC adoption averages 16.1 percent across all manufacturing firm-years and 5.1 percent across all wholesaler firm-years in our data. These averages are larger than the baseline UPC adoption rates because of the employment weights.

Finally, the data on trademark applications come from the USPTO Trademark Case Files Dataset (Graham, Hancock, Marco, and Myers, 2013). To merge these data to the Business Register, we rely on a matching procedure described in Dinlersoz, Goldschlag, Myers, and Zolas (2017). We use the trademark data to construct two outcome variables: a count of new applications in a given calendar year, and a dummy indicating that a firm applied for at least one new trademark in a particular year. Because firms can file “intent to use” applications for a TM that is never actually used in commerce, we restrict both

\textsuperscript{14}In some specifications we define $\{s(e) \setminus i(e)\}$ to be the set of establishments $j$ that share a three-digit ZIP code with establishment $e$ but are not owned by the same firm as establishment $e$. 
of our outcome measures to applications that eventually produce a registered trademark.\footnote{Though trademarks need not be registered, a federal registration affords nationwide protection and is required for enforcement in federal court. Registration implies that a firm has provided evidence that a TM was actually used in commerce. The intent-to-use application procedure became available in 1989, towards the end of our sample period, through the Trademark Law Revision Act of 1988.} To avoid double counting TMs that change hands, we restrict our counts to the original applicant.

Table 1 shows that 8.1 percent of manufacturing firm-year observations belong to firms that applied for at least one new trademark during the sample period; the corresponding figure for wholesale is 4.1 percent. The annual probability of filing was 1.6 percent for manufacturers and 0.7 percent for wholesalers, and the mean number of new applications per year was 0.035 and 0.012 respectively.

4 Diffusion

4.1 Descriptive Evidence

Dunlop and Rivkin (1997) and Dunlop (2001) document the diffusion of UPC registrations across sectors and time. They show that, through 1975, nearly two thirds of registrations were by food and beverage companies (manufacturers, wholesalers, and retailers) but that, by 1982, these firms constituted a minority of new registrations. According to Abernathy, Dunlop, Hammond, and Weil (1995), the apparel industry adopted the UPC in 1987.

Unlike these previous studies, we have data not only on UPC registrants but also on non-adopters, allowing us to document the evolution of the \textit{share} of adopters. Census data also allow us to calculate UPC registrants’ share of industry employment and revenues.

We start by partitioning all manufacturing firms, in each Economic Census year, by revenue quartile, and calculating the share of each quartile that has registered for a UPC by that year. The relative registration rates are shown on the left-hand-side panel of Figure 3. Among firms in the largest quartile, approximately 2 percent registered for a UPC by 1977,
and nearly 10 percent registered by 1992. Smaller firms have lower registration rates; no more than 2 percent of firms in the third and fourth quartiles registered for a UPC by 1992. A similar calculation for wholesale firms is shown, using the same axis, on the right. Overall registration rates in wholesale are lower, but the ranking of adoption rates by quartile is very similar. We omit retail firms from this figure because adoption rates in retail are extremely low; they follow a similar pattern, but at a much lower level.

Our data also reveal differences in UPC adoption across manufacturing industries. The UGPC was initially a grocery product code, intended for use by food manufacturers, retailers, and wholesalers. After a slow start, by 1980, Harmon and Adams (1984, p. 7) report that more than 90 percent of grocery products displayed barcodes. Soon, general-merchandise stores “noted the benefits of uniform product coding [...] and began to demand that their vendors adopt the U.P.C.” (Dunlop and Rivkin, 1997, p. 5), making it “a condition of doing business” (Martin, 2001, p. 37). Both Kmart and Wal-Mart make regular mention of the UPC in their annual reports starting around 1982, and Figure 1 shows a bunching of total registrations in 1983.

Figure 4 reinforces the idea that the UPC was widely adopted within the grocery supply chain before spreading to general merchandise. Each panel plots UPC adopters’ share of firms and employment within six selected manufacturing industries. All panels are on the same scale, but the firm share and employment share use different axes.

The UPC registration rate in food manufacturing (top left panel) is about 20 percent in

\[^{16}\text{Wholesalers may be merchant wholesalers, which are intermediaries that buy inputs and may package, repackage, or label them for sale to retailers, and manufacturers’ sales and branch offices, which act as brokers and do not take possession of the goods they sell. We cannot distinguish these two types of wholesalers in the data, but believe that the vast majority of wholesaler registrations belong to the former category. Starting in 2002, NAICS has distinguished between these two types of wholesalers; that year, merchant wholesalers accounted for 93 percent of wholesale establishments and 90 percent of wholesale revenues (U.S. Census Bureau, 2005). Ganapati (2016) uses this classification to study the role of merchant wholesalers in the supply chain. Dinlersoz, Goldschlag, Myers, and Zolas (2017) find that wholesalers are among the firms most likely to pursue trademarks.}\]

\[^{17}\text{As detailed in the Data Appendix, firms are classified by their predominant industry. Firms that operate multiple establishments may be classified in one industry despite having some establishments, perhaps a significant number, in other industries.}\]
1975, and increases to about 80 percent by 1992. The employment share, however, remains fairly stable at 60 percent: large firms registered early, and later registrants are small. In chemical production, which includes pharmaceuticals, adoption by large firms is early but both the employment share and the firm share of adopters increases steadily over time. Both food and chemical manufacturers are likely to sell through the grocery supply chain. However, the other four industries in Figure 4 (apparel, electronics, furniture and textile manufacturing) tend to supply general-merchandise retailers. For these four industries, growth in UPC adoption begins in the early 1980s and takes off more slowly, though employment growth exceeds firm growth because here, too, larger firms adopt earlier.

4.2 Network Effects

The UPC is a classic case study for two-sided network effects. The argument, as advanced by Dunlop and Rivkin (1997), is that upstream manufacturers had no incentive to make the investments — up to $10,000 for a UPC registration, plus the cost of redesigning product labels and, possibly, printing technology necessary to print precise barcodes that would not smear — until a critical mass of downstream firms had the means to take advantage of these investments. Downstream firms, meanwhile, had little incentive to make their own investments — in scanners and other computer hardware and software, and in employee training — until a critical mass of upstream firms printed barcodes on their products. Overcoming what Dunlop and Rivkin call the “chicken-and-egg” problem was the goal of the UGPC Council. The UGPCC believed that the critical mass on the manufacturing side of the market was 75 percent of grocery-product labels with a barcode, and on the retail side, 8,000 supermarkets with scanners installed (Dunlop and Rivkin, 1997, p. 28).

Estimating two-sided network effects directly is impossible with the current data. There are at least two types of hurdles: data problems and identification problems. One data
problem is that we do not have precise information on supply chains. A second data problem is that we do not have information on adoption of complementary technologies, such as barcode scanners, by downstream firms. If we had data on supply-chain links and scanner adoption, we could estimate the system of equations

\[
\Delta \text{UPC}_{it} = f \left( \mathbb{E}_t \left( \text{Scanner}_{j(i),t+1} \right) \right) \\
\Delta \text{Scanner}_{jt} = g \left( \mathbb{E}_t \left( \text{UPC}_{i(j),t+1} \right) \right)
\]

where the outcomes \( \Delta \text{UPC}_{it} \) and \( \Delta \text{Scanner}_{jt} \) indicate that firm \( i \) registered for a UPC prefix or firm \( j \) installed a scanner in year \( t \), and the explanatory variables \( \mathbb{E}_t \left( \text{Scanner}_{j(i),t+1} \right) \) and \( \mathbb{E}_t \left( \text{UPC}_{i(j),t+1} \right) \) indicate the expected future stock of scanning stores or bar-coding manufacturers within the firm’s supply chain.

The identification problem is that a simultaneity between upstream UPC registrations and downstream adoption of scanners and other complementary technology can be causal in one or both directions, or may be due to an omitted variable; without an instrument for either or both we cannot determine the relative importance of the two channels.

Instead of estimating (1) directly, we use a reduced-form approach. An implication of two-sided network effects is the benefit of adoption is higher for firms when a larger share of their competitors have also adopted the UPC, because in those settings downstream firms are more likely to have adopted the complementary technologies. To test this hypothesis, we estimate a hazard model of UPC adoption as a function of adoption by rivals in the same four-digit SIC code:

\[
\text{UPC}_{it} = \beta \text{UPC}_{j(i)t} + \lambda_{at} + \varepsilon_{it}
\]

where the outcome UPC indicates that firm \( i \) registered in year \( t \), and the main explanatory

---

18 We can infer supply chains from a coarse correspondence between product lines carried and upstream UPC adoption, but do not use those data in this draft.  
19 We have data on scanner adoption, at the store level, only for grocery stores up to 1984, from Basker (2012). 

13
variable \( \text{UPC} \) is the employment-weighted adoption rate for firms competing with firm \( i \) in the same year. The model includes a full set of firm-age by calendar-year fixed effects (\( \lambda_{it} \)), and each firm is retained in the data only until the year when it registers, so that \( \beta \) can be interpreted as the change in the hazard of UPC adoption if all of a firm’s rivals switched from non-adopters to adopters.\(^{20}\) This specification produces unbiased estimates of the spillover \( \beta \) if the variation in rival adoption is orthogonal to \( \epsilon_{it} \). This could be the case, for example, if the benefits of adoption (per SKU) are constant across manufacturers, but scale economies produce exogenous variation in costs of UPC labeling.

Estimates of this model are reported in Table 2. For both manufacturing and wholesale firms, we find statistically significant and economically meaningful positive spillovers (or strategic complementarities) in UPC adoption. These are measured by the coefficients in the first and third column of the table, for the variable labeled \( \text{UPC}_{j(i)t} \). The standard deviation of the within-industry spillover variable (as reported in Table 1) is 0.17 for manufacturers. This implies that a one standard deviation increase in rival UPC adoption more than doubles the baseline hazard of manufacturer UPC adoption (from 0.34 to 0.82 percent per year).

This result could be caused by at least two types of network-effect mechanisms: unobserved changes in downstream scanner adoption (or expectations of scanner adoption) that raise the returns to upstream UPC adoption within a supply chain, or a coordinated effort by upstream firms to start the scanner-adoption bandwagon.\(^{21}\) These results could also be caused by industry-level unobservables that increase the returns to adoption, such as a correlated reduction in costs. We find the network effects interpretation more plausible

\(^{20}\)Because the LBD starts in 1975, firm age is censored at \((t - 1975)\) for wholesalers. We are able to go back to 1972 for manufacturers that appeared in the 1972 Census of Manufactures.

\(^{21}\)There is an interesting anecdote about the role of the expectations operator in promoting UPC adoption within the grocery sector. Brown (1997, p. 116) reports “widespread misunderstanding of the 1973 economic projections by McKinsey. McKinsey had said that if there were five thousand scanning stores in 1975, then the savings attributable to scanning would be significant enough to justify the [manufacturer] investment. This was widely interpreted to mean that McKinsey had predicted there would be five thousand stores by 1975, over 4900 more than were in place.” He does not speculate as to whether the misunderstanding was an accident.
because the costs of adoption were small, and do not seem to have a large industry-specific component.

To reinforce the idea that these estimates measure network effects, we estimate a placebo model that seeks evidence of geographic spillovers. If strategic complementarities in UPC adoption are caused by downstream scanner adoption, then we would expect to see much smaller spillovers from geographically proximate firms that sell through different supply chains. To perform the placebo test, we re-define \( \overline{UPC}_{j(i)t} \) as the average UPC adoption rate by three-digit zipcodes (again building up from establishment-level location data), and re-estimate Equation (3). The coefficients in the second and fourth columns of Table 2 show that in manufacturing, the geographic spillovers are an order of magnitude smaller than the industry spillovers (and not statistically significant), whereas geographic spillovers in wholesaling are roughly 20-percent of the intra-industry spillovers.

4.3 Retail Registrations

Although a UPC registration is not required for implementation of scanning, hundreds of retail firms did register for a UPC. These firms are disproportionately large: compared to Figure 3, which shows higher registration rates for manufacturing and wholesale firms with sales above the median than for firms with sales below the median, the majority of retailers that registered for a UPC by 1992 were in the top quartile of the sales distribution. As in manufacturing and wholesaling, food retailers registered much earlier than apparel retailers, a distinction that is magnified in employment-weighted terms. (These figures are omitted to comply with Census rules on disclosure avoidance.)

If registration was not needed to implement scanning and complementary technologies such as EDI, why did retail firms register for UPCs? Some retail firms operate one or more upstream establishment, such as a supermarket chain that has one production facility for private-label items, and these retailers may have registered in order to implement barcoding on their private labels. Others, particularly large retailers that had the most to gain from
implementing scanning in their stores, may have registered to show support for the system. Finally, grocery stores with meat and deli departments that wanted to print their own barcodes on variable-weight items would also have needed to register for a UPC.\(^{22}\)

We explore the relative importance of these three explanations by estimating, for the set of retail firms,

\[
\text{UPC}_{it} = \beta \ln (\text{Establishments}_{it}) + \phi \text{Vertical}_{it} + \gamma \text{Grocery}_{it} + \lambda_{at} + \varepsilon_{it} \tag{4}
\]

where, as before, UPC\(_{it}\) indicates that firm \(i\) registered in year \(t\), and each firm is retained in the data only until the year when it registers. The explanatory variable Establishments is the number of establishments (stores, warehouses, offices, etc.) that firm \(i\) operates in year \(t\); Vertical is an indicator for the firm having one or more establishments in the wholesale and/or manufacturing sectors; Grocery is an indicator for the firm having one or more establishments with three-digit SIC 541; and \(\lambda_{at}\) is a separate calendar-year effect for each firm age \(a\). SIC 541 includes supermarkets (which are more likely to have meat and deli department) but also convenience stores. Standard errors are clustered by firm SIC.

Our estimate of \(\beta\) is positive, economically meaningful, and statistically significant: larger retail firms are more likely to register for a UPC than smaller ones, consistent with the notion that larger retailers had a higher value of signaling a commitment to the new technology even prior to a large-scale installation of scanners. We also find that \(\phi\) is positive, economically large, and statistically significant: even controlling for the total number of establishments in the firm, vertically integrated firms were more likely to register for a UPC, possibly for their private labels. The estimate of \(\gamma\), however, is near zero and statistically insignificant, so either the variable Grocery does not capture the “supermarket effect” because of the large number of convenience stores and smaller grocery stores included in SIC 541, or

\(^{22}\)Early in the UPC implementation process, Toledo Scale and Hobart Scale, two manufacturers of meat scales, worked with IBM to ensure that in-store meat scales would be able to print labels with UPC symbols (Selmeier, 2008, pp. 131–132).
the value of printing barcodes in store was not sufficiently strong as to induce registration. This latter point may be partly because the firm knows its own rate of scanner adoption: there is no expectation operation necessary for retailers in Equation (1), so actual (low) scanner adoption rates determine vertically integrated retailers’ UPC registrations.

5 Firm-Level Impacts of UPC Adoption

5.1 Employment

We measure the impacts of UPC adoption on manufacturing and wholesale firms by estimating a difference-in-difference regression:

\[
\ln(\text{Employment}_{it}) = \alpha_i + \lambda_{at} + \beta \text{UPC}_{it} + \varepsilon_{it} \tag{5}
\]

where \( \text{Employment}_{it} \) is firm \( i \)'s employment in year \( t \), aggregated from establishment-level data in the LBD; \( \alpha_i \) is a firm fixed effect; and \( \text{UPC}_{it} \) is an indicator that turns on for a firm that registered for a UPC by year \( t \). The year-age fixed effects capture many unobservable factors, including the fact that firms tend to grow as they age, and that young and old businesses react differently to business-cycle shocks (Haltiwanger, Jarmin, and Miranda, 2013; Fort, Haltiwanger, Jarmin, and Miranda, 2013). Standard errors are clustered by four-digit firm SIC to allow for arbitrary autocorrelation in the error term \( \varepsilon_{it} \) as well as arbitrary correlation across firms in the same industry.

We estimate this equation separately for firms in the manufacturing and wholesale sectors. In each case, we estimate two specifications. The first specification uses all non-adopting firms in the sector as controls. The second specification uses a matching estimator. Each firm in the matched sample has a unique match, chosen as follows. First, we identify the pool of potential matches for firm \( i \), which registered for a UPC in year \( t \), as firms that had nonzero employment in year \( t \) and did not match to the UPC files by 1992. If firm \( i \)
is observed for the first time in the year of registration, we randomly assign one firm of the same vintage in the same year as a match. If firm $i$ is observed for the first time one year prior to registration, we assign a match using its age and vintage and year $(t - 1)$ employment level.\(^{23}\) For firms ages two through four at registration, we match using vintage, year $(t - 1)$ employment, and log employment growth between year of birth and year $(t - 1)$.\(^ {24}\) Finally, we match firms aged five and over at the time of registration to other firms that are at least five years old in year $t$ by year $(t - 1)$ employment and by log employment growth between year $(t - 5)$ and year $(t - 1)$. Registrants that do not have a matched control firm are dropped from this sample.\(^ {25}\)

The coefficient estimates are reported in Table 3. All of the OLS and the matching coefficients are positive and statistically significant. However, the matching estimators yield smaller point estimates, consistent with selection effects. That is, UPC adopters grew faster than non-adopters before they registered for a UPC. The baseline OLS estimates for manufacturing imply a 16 percent increase in employment following UPC adoption, whereas the matched sample estimates suggest a nine percent increases. For wholesaling, the baseline OLS estimate suggests a 23 percent increase in employment, compared to a 14 percent increase for the matched sample.

Even with matching, it is hard to say to what extent these regressions estimate a selection effect as opposed to a causal effect of UPC adoption. To get a better handle on this question, we estimate a series of event-study regressions of manufacturing establishment employment

\(^{23}\)We bin employment in 50 bins per year, each with two percent of the firms. We drop any bins whose maximum size exceeds 110 percent of their minimum size to ensure that employment at matched control observations is within 10 percent of employment at treated observations.

\(^{24}\)We find the closest match on employment growth, with the restriction that the two firms’ employment growth cannot differ by more than 0.5 percent.

\(^{25}\)Matching firms always share a predominant sector (i.e., manufacturing or wholesale) but we do not restrict matching firms to be in the same industry. We made this decision for two reasons. First, as a practical matter, restricting to the same four-digit or even two-digit industry reduces the number of possible matches for each treated observation, and would therefore decrease the quality of the match or the number of matches, or both. Second, on a more conceptual level, within-industry non-adopting firms may not be a valid control group because they can be affected by competitors’ adoption of the UPC.
on UPC adoption. Our main specification is:

\[
\ln(\text{Employment}_{it}) = \alpha_i + \lambda_{at} + \sum_k \beta_k \Delta\text{UPC}_{i,t+k} + \varepsilon_{it}
\]  (6)

where \(\alpha_i\) and \(\lambda_{at}\) are as above, but now \(\Delta\text{UPC}_{i,t+k}\) is a series of indicators that turn on only if the firm registered for a UPC in year \((t+k)\), and \(\beta_k\) are a series of age-relative-to-adoption coefficients. We normalize \(\beta_{t-1} = 0\). The variable \(\Delta\text{UPC}_{i,t-6}\) is turned on if the firm adopted a UPC registration six or more years in the future. To ensure that we do not include future adopters in the control group, we restrict this regression to observations in 1986 and prior years.

Figure 5 shows the event-study coefficients for the full sample. The connected dots correspond to point estimates, and the error bars are upper and lower 95 percent confidence limits. We observe a very strong selection effect for both manufacturing and wholesale firms: in the years prior to registration, soon-to-adopt firms grow much faster than controls. Once they register, however, UPC-adopting firms stop growing disproportionately, at least for a few years. The selection effects revealed in this event study regression are consistent with our observation that firm size and UPC adoption are strongly correlated, since faster pre-adoption growth naturally produces a larger firm.

Figure 6 shows the corresponding matched sample results. By construction, relative employment of adopters and non-adopters at \((t-5)\) and, to a lesser extent, from \((t-4)\) to \((t-2)\), is nearly identical and statistically indistinguishable. (However, adopting firms that were at least six years old at the time of adoption do show more employment growth than their matched controls between years \((t-6)\) and \((t-5)\).) Following adoption, the treated and control firms clearly diverge: employment at adopting firms increases for a couple of years, and plateaus at a new, higher level. The level of the coefficients from \((t+1)\) to \((t+11)\) in the matched specification is in the neighborhood of the difference-in-difference matching coefficient.
The abrupt increase in relative employment observed in these matched difference-in-difference models is a striking result. If UPC registration were a proxy for the adoption of some broader set of technological and organizational changes, we would expect a slow-and-steady divergence in employment. The discrete jump suggests to us that manufacturers were adopting the UPC to integrate with retail supply chains. This does not mean we believe that UPC adoption caused retail orders to arrive — it seems equally likely that anticipated demand shocks caused a firm to adopt UPC. However, these results do suggest that UPC adoption was a necessary condition for achieving scale through partnering with larger downstream firms.

As a robustness check, we also examined the link between UPC adoption and firm revenue. Because revenues are available only in five-year intervals, we cannot estimate an event-study specification, but we do estimate fixed effects specifications, replacing the LHS variable in Equation (5) with log revenue. Given the pattern of post-adoption dynamics in Figures 5 and 6, the once-and-for-all assumption implicit in the difference-in-difference specification seems reasonable. As with employment, we estimate this regression separately for manufacturing and wholesale, for both the full and matched samples. The coefficient estimates when using log revenue as the outcome variable are very similar to the ones reported in Table 3.

5.2 Network Effects

In this sub-section, we extend our prior analysis of network effects by testing whether the impact of UPC adoption on employment is greater when more of a firm’s rivals have also registered. In this specification, rivals’ registrations serve as a proxy for downstream complementary technology adoption, or the expectation thereof. Although downstream UPC-readiness should increase the benefits of UPC adoption for a focal firm, rival adoption could also reduce those benefits by removing opportunities for differentiation. So, in practice, we
estimate the net effect of these two forces, which we expect to provide a lower bound for the casual impact of downstream complementary technology adoption.

To implement our test for network effects, we add two variables to the difference-in-difference specification in Equation (5): the employment-weighted UPC adoption rate for each firm’s competitors, \( \overline{\text{UPC}} \) (by four-digit industry), and an interaction of the competitors’ adoption intensity with the firm’s own adoption indicator. We estimate

\[
\ln(\text{Employment}_{it}) = \alpha_i + \lambda_{at} + \beta \overline{\text{UPC}}_{it} + \gamma \overline{\text{UPC}}_{j(i)t} + \delta \overline{\text{UPC}}_{it} \overline{\text{UPC}}_{j(i)t} + \varepsilon_{it}
\]  

(7)

The coefficients on the direct effect, \( \beta \), and on the interaction term, \( \delta \), for each of the four samples (OLS and matched, manufacturing and wholesale) are shown in Table 4. The main effect of UPC adoption (or the impact for a firm in an industry with no other adopters, \( \beta \)) is reported in the first row of the table. The direct effect of adoption is positive and statistically significant in all four models. The interaction term, which we interpret as a measure of network effects, is positive and statistically significant in three of the four cases (the exception being the full sample for manufacturing). In terms of magnitudes, if we focus on the matched sample results, a one standard-deviation increase in rival UPC adoption raises the marginal effect of own-UPC adoption on employment from 4.6 to 6.7 percent for manufacturers, and from 10.4 to 11.9 percent for wholesalers.\(^{26}\)

One concern with these estimates is that, as explained earlier, we do not match firms on industry or industry characteristics, in part because we are concerned that within-industry non-adopting firms may not be a valid control group because they can be negatively affected by competitors’ adoption of the UPC. As a result, even in the matched sample, the coefficient \( \delta \) is estimated off differences in adoption rates across industries. As a robustness check, we have re-estimated these models on a matched sample within a four-digit firm SIC. This

\(^{26}\)The marginal effect of \( \text{UPC}_{it} \) is equal to \( \beta + \delta \overline{\text{UPC}}_{j(i)t} \). For manufacturers we estimate \( \delta = 0.123 \) and (from Table 1) the standard deviation of \( \overline{\text{UPC}}_{j(i)t} \) is 0.17, so a one SD shift in \( \overline{\text{UPC}}_{j(i)t} \) increases the predicted marginal effect by \( 0.17 \cdot 0.123 = 0.021 \).
reduces the sample size by 40 percent. Our coefficient estimates change very little for this sample, but the estimated interaction effect, $δ$, loses statistical significance due to larger standard errors.

### 5.3 Survival

One issue with these difference-in-difference regressions is that survival rates may differ for adopters and non-adopters. In particular, firms that do not anticipate surviving for long may avoid the costs of UPC adoption. Because we use log employment as the LHS variable in the employment regressions, non-survivors drop out, possibly biasing downwards our estimates of the impact of adoption. To see how important this effect is, we estimate the following hazard model for the full sample of firms in each sector,

$$\text{Exit}_{it} = \lambda_{at} + \beta \text{UPC}_{it} + \varepsilon_{it}$$  (8)

where $\text{Exit}$ is an indicator that equals one if and only if firm $i$ is observed for the last time in year $t$, and $\text{UPC}$ is, as in Equation (5), an indicator that turns on if the firm has adopted by year $t$.\(^{27}\)

We also estimate a variant of this hazard regression using the matched samples. For these, we estimate

$$\text{Exit}_{it} = \lambda_{at} + \tau \text{Registrant}_{i} + \pi \text{PostRegistration}_{it} + \beta \text{UPC}_{it} + \varepsilon_{it}$$  (9)

where $\text{Registrant}$ is an indicator that turns on if firm $i$ registered for a UPC by the end of 1992; $\text{PostRegistration}$ is an indicator that turns on for both UPC adopting firms and their matched controls in the year of adoption, and remains on thereafter; and $\text{UPC}$ is the interaction of $\text{Registrant}$ and $\text{PostRegistration}$. The coefficient $\tau$ on the treatment indicator tells us whether

\(^{27}\)This sample ends in 1991, because our last observation of all firms is 1992.
overall exit rates differ across the treatment and control groups. The coefficient \( \pi \) on the post-treatment indicator tells us whether the pair of treatment and control firms is, on average, more likely to exit later than earlier. By construction this coefficient must be positive, because firms cannot exit before they (or their treated counterparts) register for a UPC. Finally, the coefficient \( \beta \) on the interaction term indicates whether, and to what extent, exit rates are lower among adopters than their matched control firms after the adoption date.

We estimate hazard models for both the full and matched samples, and report the results in Table 5. From Table 1, manufacturing firms in our panel have a 9.7 percent probability of exit, and wholesalers have a 9.6 percent probability of exit. Our hazard model estimates imply that this exit rate declines by around two percentage points for manufacturing firms and around 1.5 percentage points for wholesalers. The estimates are quite stable across the full sample and matched sample models. To summarize, UPC registration is associated with a lower exit rate, even relative to matched control firms that have similar prior size and growth rates.

5.4 Trademarks

Our final set of analyses examine the link between UPC adoption and innovation. Several scholars have suggested that as UPCs lowered the cost of tracking and managing inventory, retailers became willing to stock a greater variety of products, which in turn increased the incentive for manufacturers to experiment with new product varieties. For instance, Dunlop (2001, p. 20) writes that, “The diffusion throughout the Food and Beverage sector has been steady with associated product proliferation, much larger stores and the addition of numerous new departments and an approach to the early objective of one-stop shopping.” As a proxy for variety-increasing product innovation, we count the number of trademark (TM) applications that resulted in a TM registration issued to the focal firm.

Figure 7 provides some time-series evidence to help motivate our analysis. Two of the
lines in this graph are ocular reproductions of data reported in Sullivan (1997).\textsuperscript{28} The solid line shows a time series of new product introductions in the grocery sector collected by the periodical \textit{New Product News}, and the dashed line shows the average number of SKUs per grocery store reported by \textit{Progressive Grocer}.\textsuperscript{29} The dotted line is a count of new TM applications for grocery-related products that we constructed from the USPTO data.\textsuperscript{30,31}

Figure 7 helps motivate our firm-level analysis in two ways. First, it shows that TM applications are strongly correlated with new product introductions and the expansion in SKUs on retail shelves. This suggests that it is reasonable to use TM applications as a proxy for variety-expanding product innovation. Second, it shows that all three time-series appear to have a trend-break sometime between 1980 and 1985, which is roughly the time-period when the UPC was diffusing through the grocery supply chain (see Figures 1 and 4).

Our empirical specification is the same two-way fixed effects model presented in Equation (5), except that for the outcome variable, we replace log employment with either a count of new TM applications filed by firm $i$ in year $t$ or an indicator that a firm filed for one or TMs. Receiving a TM is not a common event. Only approximately eight percent of the observations in our manufacturing sample, and four percent of the observations among wholesalers, come from firms that received one or more TMs (Table 1). From this, we can infer that conditional on ever receiving a new TM, manufacturers average 0.43 TM applications per year, and wholesalers average 0.29 TM applications per year.

\textsuperscript{28}We resorted to the eyeball method because her original data have been lost.

\textsuperscript{29}The latter series has a gap in coverage between 1972 and 1982, as shown in the figure.

\textsuperscript{30}In order to restrict our count of TM applications to the grocery industry, we focus on applications with a 3-digit Nice code corresponding to food, beverages, pharmaceuticals or paper products. For applications that match to more than one Nice code, we randomly selected a single code for the analysis.

\textsuperscript{31}The USPTO Trademark Case Files data show a sharp increase in applications around 1977, because TMs registered prior to 1977 and later abandoned were not recorded in the USPTO computer systems (Graham, Hancock, Marco, and Myers, 2013, p. 32). The trend in registrations is essentially flat from 1963 to 1974, and again from 1977 to 1981, but exhibits a break in 1975 and 1976. We do two things to adjust for this in the figure. First, we inflate values for 1973 and prior years by multiplying the actual TM count by the ratio of 1977 to 1974 TMs. Second, we linearly interpolate 1975 and 1976 values from 1974 and 1977 values.
The results of our trademark analysis are presented in Table 6. The difference-in-difference estimates for both manufacturing and wholesale, first using the TM indicator and then the TM count as outcomes, show a large and statistically significant increase in trademarking following UPC registration. For example, the coefficient in the first column of the table indicates that the probability that a manufacturer files for a new TM in a given year increases by 230 percent following UPC adoption.32 In the second column, we observe a six-fold increase in the count of new TM filings after UPC adoption.

We conducted a variety of robustness checks. Although the results are not reported because of disclosure considerations, they are all broadly similar to the results in Table 6. Our first set of robustness checks focused on the outcome variable. Instead of counting all TM filings that led to a registration, and dating them to the filing year, we created a variable that counts applications for which there is data on the date of first use in commerce, linking each new TM to that date. Our second set of robustness checks focused on the estimation sample. In particular, we re-estimated each model keeping only the firms that actually applied for at least one TM during our sample period.33 This changes the magnitude of the coefficients, but the size of the marginal effect remains similar. Finally, we estimated a Poisson specification of (5) to account for the skewed outcome and many zeros. Results were similar. We were not able to create a matched sample analysis for TMs, as we did with employment, because of the rare-event nature of the outcome variable.

6 Concluding Remarks

Bar codes were a key component in a broad set of innovations that dramatically lowered the cost of managing inventory in retail supply chains. Scholars have suggested that this had far-reaching implications, including the rise of the new “bog box” format (e.g., Holmes,

32This marginal effect is calculated relative to the mean of the outcome, i.e., (0.053/0.016 − 1).
33In principle, this might select on outcomes.
2001, and Dunlop, 2001). This paper is the first to analyze the upstream and network effects of UPC adoption, providing a natural complement to the analysis of retail scanner adoption in Basker (2012) and a new addition to the empirical literature on the effects of adopting industry standards.

We show that UPC adoption is strongly correlated with firm size and that the timing of UPC adoption varied across industries. Many large food-and-drug manufacturers had already adopted the UPC by the mid 1970s, whereas adoption by apparel, furniture and textile manufacturers remained at low levels into the early 1980s. This pattern is consistent with the idea that upstream UPC adoption was driven by (expectation of) downstream installation of complementary scanning technology, which began in the grocery sector and was later implemented by large general merchandisers. We provide additional evidence on this point by estimating a reduced-form model of within-industry spillovers in UPC adoption, and find positive effects for both manufacturers and wholesalers.

Our investigation of the impacts of UPC adoption suggests that upstream firms benefited from early adoption on several margins. First, survival rates were higher for adopters post-adoption than for their matched controls. Second, both revenue and employment increased concurrent with adoption, and stayed higher, consistent with receiving larger orders from retailers. Finally, we find evidence consistent with the presence of positive network effects: both the rate of UPC adoption and the increase in employment following adoption are larger when more of a firm’s competitors have already adopted. The likely mechanism is that the more intensively adopting industries have a downstream customer base that has installed scanners and implemented other processes, such as EDI, complementary with UPCs. These downstream firms, therefore, have a preference for placing orders with suppliers that themselves have UPC capabilities.

The timing of the employment effects revealed by our event study regressions suggests that UPC adoption is associated with business diversion, whereby manufacturers integrate into the supply chain of large downstream retailers, rather than the effect of adopting a
suite of UPC-related technologies and business practices, which would yield more gradual and consistent increases in employment. Of course, this interpretation immediately leads to questions about selection versus causation, because it seems unlikely that UPC adoption would be sufficient to generate a large increase in demand, even if it is a necessary condition for working with large downstream retailers. It is important to recognize that our preferred matched-sample difference-in-difference specification does not remove these selection effects. By construction, it does eliminate any differential pre-adoption trend, but it assumes that, within a matched pair of firms, assignment of a UPC registration is random, an assumption that almost certainly fails to hold in the data. This is probably most obvious on the survival margin: a firm that anticipates that it may need to shut down in the near future, for example because of limited financing or low demand, is much less likely to adopt new technology than a firm that expects to remain in operation for years.

The downstream effects of UPC adoption are harder to assess. Because we do not have explicit data on supply chains — in other words, which retailers purchase from which suppliers — we cannot test directly the hypothesis that upstream UPC adoption increases downstream store size or selection. Time-series evidence, however, supports the idea that the retail sector responded to the UPC by increasing store assortment. For example, we show that the rate of growth in unique products (SKUs) stocked by in supermarkets, the number of new product introductions in the grocery sector, and the number of new trademark applications filed by food and grocery manufacturers all increased dramatically starting in the early 1980s, as barcodes and scanners became pervasive in that industry. Moreover, by linking UPC adoption to firm-level information on trademarking from Dinlersoz, Goldschlag, Myers, and Zolas (2017), we show that the increase in trademarking occurs within firms that registered for UPC, suggesting a direct link between improved supply-chain coordination and increased new-product variety.

Overall, we see the UPC as a stand-in for a combination of demand effects, management practices, and the adoption of other technologies. Adding bar codes does not in itself “cause”
a manufacturing or wholesale firm to grow. At most, it may divert demand from competitors if the firm’s downstream customers value having a UPC on its products. However, the correlations between UPC adoption, growth, and innovation suggest that UPC-adopting firms may also be ones that are more likely to adopt other new technologies, use innovative management practices, and seek new markets.
References


Figure 1. UPC Registrations

(a) By Year

(b) By Size Class

Figure 2. UPC-BR Match Rate

(a) By Year

(b) By Size Class
Figure 3. UPC Diffusion by Firm Revenues

Figure 4. UPC Diffusion for Selected Two-Digit SIC Manufacturing Industries
Figure 5. Event Study Coefficients: Manufacturing and Wholesale Employment

(a) Manufacturing
(b) Wholesale
Note: Vertical bars represent 95 percent confidence intervals

Figure 6. Event Study Coefficients: Manufacturing and Wholesale Employment, Matched Samples

(a) Manufacturing
(b) Wholesale
Note: Vertical bars represent 95 percent confidence intervals
Figure 7. New Products Introductions, U.S. Trademark Applications, and Average Stock-Keeping Units per Store in the Grocery Sector

Note: TM Applications before 1977 adjusted due to missing data (see text for details).
Table 1. Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Manufacturing Mean</th>
<th>Manufacturing SD</th>
<th>Wholesale Mean</th>
<th>Wholesale SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employees</td>
<td>72.8</td>
<td>1,540</td>
<td>14.9</td>
<td>193</td>
</tr>
<tr>
<td>[\text{EverUPC}]</td>
<td>0.038</td>
<td>0.190</td>
<td>0.022</td>
<td>0.145</td>
</tr>
<tr>
<td>[\text{UPC}]</td>
<td>0.019</td>
<td>0.135</td>
<td>0.010</td>
<td>0.100</td>
</tr>
<tr>
<td>Rival UPC Share(^a)</td>
<td>0.161</td>
<td>0.170</td>
<td>0.051</td>
<td>0.073</td>
</tr>
<tr>
<td>[\text{EverTM}]</td>
<td>0.081</td>
<td>0.266</td>
<td>0.041</td>
<td>0.195</td>
</tr>
<tr>
<td>[\text{TM}]</td>
<td>0.016</td>
<td>0.117</td>
<td>0.007</td>
<td>0.081</td>
</tr>
<tr>
<td>TM count</td>
<td>0.035</td>
<td>0.614</td>
<td>0.012</td>
<td>0.212</td>
</tr>
<tr>
<td>[\text{Exit alive at } t-1]</td>
<td>0.097</td>
<td>0.296</td>
<td>0.096</td>
<td>0.295</td>
</tr>
<tr>
<td>Firms(^b)</td>
<td>779,300</td>
<td></td>
<td>866,500</td>
<td></td>
</tr>
<tr>
<td>Observations(^b)</td>
<td>5,112,400</td>
<td></td>
<td>5,621,800</td>
<td></td>
</tr>
</tbody>
</table>

Notes: \(^a\)Employment-weighted share of firm’s competitors, by four-digit SIC, that have registered for a UPC.
\(^b\)An observation is a firm-year. Observation counts rounded to the nearest hundred to comply with Census rules on disclosure avoidance.

Table 2. UPC Adoption: Industry vs. Geographic Spillovers

<table>
<thead>
<tr>
<th>Spillover</th>
<th>Manufacturing SIC</th>
<th>Manufacturing ZIP</th>
<th>Wholesale SIC</th>
<th>Wholesale ZIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>UPC(_{(i),t})</td>
<td>0.0282</td>
<td>0.0005</td>
<td>0.0277</td>
<td>0.0050</td>
</tr>
<tr>
<td></td>
<td>(0.0022)**</td>
<td>(0.0003)</td>
<td>(0.0062)**</td>
<td>(0.0003)**</td>
</tr>
</tbody>
</table>

Mean outcome 0.0034 0.0023
Observations\(^a\) 5,033,100 5,577,100

Notes: Outcome: UPC Adoption. Firms remain in sample until year of first UPC adoption.
\(* p<10\%; ** p<5\%; *** p<1\%
Robust SEs clustered by four-digit firm SIC or three-digit firm ZIP in parentheses. All regressions include age×year fixed effects.
\(^a\)Observation counts rounded to the nearest hundred to comply with Census rules on disclosure avoidance.
Table 3. Difference-in-Difference Regressions: Log Employment

<table>
<thead>
<tr>
<th></th>
<th>Manufacturing</th>
<th>Wholesale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full Matched</td>
<td></td>
</tr>
<tr>
<td>UPC$_{it}$</td>
<td>0.161</td>
<td>0.228</td>
</tr>
<tr>
<td>(0.012)***</td>
<td>(0.014)***</td>
<td>(0.010)***</td>
</tr>
<tr>
<td>Observations$^a$</td>
<td>5,112,400</td>
<td>5,621,800</td>
</tr>
</tbody>
</table>
| Notes:         | * p<10%; ** p<5%; *** p<1%. Robust SEs clustered by four-digit firm SIC in parentheses. All regressions include firm and age×year fixed effects. $^a$Observation counts rounded to the nearest hundred to comply with Census rules on disclosure avoidance.

Table 4. Heterogeneity by Industry UPC Intensity: Log Employment

<table>
<thead>
<tr>
<th></th>
<th>Manufacturing</th>
<th>Wholesale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full Matched</td>
<td></td>
</tr>
<tr>
<td>UPC$_{it}$</td>
<td>0.127</td>
<td>0.190</td>
</tr>
<tr>
<td>(0.023)***</td>
<td>(0.018)***</td>
<td>(0.016)***</td>
</tr>
<tr>
<td>UPC$<em>{(j(i))t}$ · UPC$</em>{it}$</td>
<td>0.091</td>
<td>0.244</td>
</tr>
<tr>
<td>(0.056)</td>
<td>(0.107)**</td>
<td>(0.085)**</td>
</tr>
<tr>
<td>Observations$^a$</td>
<td>5,112,400</td>
<td>5,621,800</td>
</tr>
</tbody>
</table>
| Notes:         | * p<10%; ** p<5%; *** p<1%. Robust SEs clustered by four-digit firm SIC in parentheses. UPC$_{(j(i))t}$ is the adoption rate at the firm’s competitors, by SIC All regressions include UPC$_{(j(i))t}$, firm and age×year fixed effects. $^a$Observation counts rounded to the nearest hundred to comply with Census rules on disclosure avoidance.
Table 5. Exit Regressions

<table>
<thead>
<tr>
<th></th>
<th>Manufacturing</th>
<th>Wholesale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full</td>
<td>Matched</td>
</tr>
<tr>
<td>UPC(_{it})</td>
<td>-0.023</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td>(0.003)***</td>
<td>(0.004)***</td>
</tr>
<tr>
<td>Ever UPC(_i)</td>
<td>0.000</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Post UPC(_{it})</td>
<td>0.084</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td>(0.004)***</td>
<td></td>
</tr>
</tbody>
</table>

Observations\(^a\) 4,800,000 207,600 5,262,900 153,200

Notes: * p<10%; ** p<5%; *** p<1%. Robust SEs clustered by four-digit firm SIC in parentheses.
All regressions include age×year fixed effects.
\(^a\)Observation counts rounded to the nearest hundred to comply with Census rules on disclosure avoidance.

Table 6. Difference-in-Difference Regressions: Trademarking

<table>
<thead>
<tr>
<th></th>
<th>Manufacturing</th>
<th>Wholesale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Any TM</td>
<td>TM Count</td>
</tr>
<tr>
<td>UPC(_{it})</td>
<td>0.053</td>
<td>0.211</td>
</tr>
<tr>
<td></td>
<td>(0.004)***</td>
<td>(0.025)***</td>
</tr>
<tr>
<td>Mean Outcome</td>
<td>0.016</td>
<td>0.035</td>
</tr>
</tbody>
</table>

Observations\(^a\) 5,112,400 5,621,800

Notes: * p<10%; ** p<5%; *** p<1%. Robust SEs clustered by four-digit firm SIC in parentheses.
All regressions include firm and age×year fixed effects.
\(^a\)Observation counts rounded to the nearest hundred to comply with Census rules on disclosure avoidance.
Data Appendix

Matching UPC Registration to Census Establishment


The Economic Census is quinquennial (every five year) and covers, with few exceptions, all business establishments with paid employees in the United States. The Census defines a business establishment as a location of economic activity and employment. In our sample, an establishment may be a manufacturing plant, a distribution center, a warehouse, a store, or an administrative office (such as a sales office or firm headquarters).

From the Economic Census, we draw all business establishments surveyed in the Census of Manufactures (CM), Census of Wholesale (CW), or Census of Retail Trade (CRT). We also draw all business establishments in the LBD in each of these three sectors. For each of the establishments, using the firm identifier in these files, we also identify any “sibling” establishments belonging to the same firms; for example, the sales office or headquarters of a firm that has one or more establishments classified as retail, wholesale, or manufacturing. We extract the names and addresses of all establishments in the relevant set of firms for each year from 1975 through 2000 from the Business Register (BR).

We match the UPC registrations to business names and addresses in the BR. Each registration in the Dunlop UPC registration file contains up to four company names (name, altname, parent company name, and division name) and up to two addresses. The BR contains up to two names (name1 and name2) and up to two addresses (mailing address and

---

34 Exceptions include most government-owned or -operated establishments, establishments operated by religious organizations, and agricultural establishments.
35 Almost all establishments in these sectors appear in the Economic Census files. However, because the Economic Census takes place only every five years, establishments that entered and exited between these years cannot be included in the Economic Census.
36 We match registrations from 1974 and prior years to the 1975 BR.
physical address). We cross match all the names and all the addresses available in any given record, by year, including prior and future BR addresses for establishments in existence in the year of a given UPC registration. We require that at least the state and either the city name or the zip code match perfectly across the two files. For records that do not have a perfect match on both name and address, we use the Levenshtein distance (edit distance) to determine how similar the names and addresses are. When one address is a street address and another is a Post Office Box, we rely heavily on the name match. We penalize names and addresses that are very common (e.g., businesses with “generic” names that use common words like American, Food, or Systems; and addresses in industrial parks or large office buildings if they are shared by a large number of tenants) and upgrade matches on names that share a unique or rare element, such as an unusual spelling of a word.

**Aggregation to Firm Level**

Most firms operate just one establishment each, and in those cases constructing firm-level variables is trivial. For multi-unit firms, we define a new UPC registration in year \( t \) as a firm that did not have a registration in year \( (t - 1) \) and that has one or more establishments match to the UPC dataset in year \( t \). That firm is assumed to have the UPC registration as long as it continues operating, whether or not the specific establishment(s) that were identified in the match continue to operate and whether or not they are divested from the firm. If a firm identifier disappears from the data, but one or more of its establishments survives and all surviving establishments have a common firm identifier in the following year, the new firm identifier is assumed to have inherited the UPC registration.

Census classifies an industry for each establishment based on the primary activity at that establishment. For example, a warehouse and a production plant belonging to the same firm would have different Standard Industrial Classification (SIC) codes. (Although the Census switched to the North American Industrial Classification System (NAICS) starting in 1997, we end our data in 1992, so we rely on SIC codes.) To aggregate the SIC to the firm level,
we use the firm’s payroll (deflated by the CPI) to determine the predominant industry in which the firm operates over the entire span in which it appears in our data.

Firm-level outcome variables available at annual frequency, by aggregating the LBD, are employment and the number of establishments the firm operates. At five-year intervals (in the Economic Census years 1977, 1982, 1987, and 1992) we also have establishment-level revenue, which we aggregate to the firm level. Firm revenues include revenues reported in the CM, CW, and CRT only.