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

Emek Basker
U.S. Census Bureau

Tim Simcoe
Boston University & NBER

December 2017

Abstract

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 Universal Product Code (UPC). We find evidence of network effects in the diffusion process. Matched-sample difference-in-difference estimates show that employment and trademark registrations increase following UPC adoption by manufacturers or wholesalers. Industry-level imports also increase with domestic UPC adoption. Our findings suggest that barcodes, scanning, and related technologies helped stimulate variety-enhancing product innovation and encourage the growth of international retail supply chains.

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, James Bessen, David Brown, James Conley, Emin Dinlersoz, Lucia Foster, Nathan Goldschlag, Michał Grajek, Dan Gross, Fariha Kamal, Alex Krasnikov, Mark Kutzbach, Florin Maican, Paul Messinger, Guy Michaels, Matilda Orth, Pham Hoang Van, Jennifer Poole, Horst Raff, Rich Richardson, Martha Stinson, Mary Sullivan, Dan Trefler, Kirk White, and seminar participants at the U.S. Census Bureau, MINES-ParisTech, Seoul National University, Cornell, the 2017 AEA (Chicago), the 2017 IIOC (Boston), the 2017 FSRDC Conference (Los Angeles), and the 2017 CAED Conference (Seoul) 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 promoting the rise of large retail chains (Sieling, Friedman, and Dumas, 2001; Foster, Haltiwanger, and Krizan, 2002). Tim Harford even named the barcode one of "50 Things That Shaped the Modern Economy."[1]

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 provide new evidence on the diffusion and impacts of the UPC by linking archival data on UPC registrations to firm-level data on employment, revenue, and trademark applications, as well as industry-level data on trade flows. Our findings illustrate the role of network effects in the adoption of the UPC system, and suggest that barcodes, scanning, and related technologies helped stimulate variety-enhancing product innovation and encourage the growth of international retail supply chains.

Previous accounts of UPC diffusion have emphasized that barcodes 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, Ham-

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

mond, and Weil, 1999; Basker, 2012). Although we lack comprehensive data on scanner installation, an indirect test of network effects is whether firms are more likely to register for a UPC when their rivals also register. We find strong evidence of within-industry spillovers in manufacturer and wholesaler UPC registration. On the retail side, we construct an establishment-level measure of upstream UPC adoption, based on the sales-weighted UPC adoption rate in different merchandise categories, and show that the hazard of grocery-store scanner adoption increases with upstream adoption.

After examining UPC diffusion, we study firm-level impacts of UPC adoption on employment, revenue, and trademarking, as well as the industry-level relationship between UPC adoption and international trade. Difference-in-difference regressions on a matched sample of UPC adopters and non-adopters show an immediate 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 promotes growth through business diversion because 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 diversion by manufacturers and wholesalers that can integrate into large retailers' supply chains more readily after registering for a UPC. The increase in employment following UPC registration is increasing in the share of the firm's competitors that have also registered, consistent with the presence of network effects.

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 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.[2] Difference-in-difference regressions show that manufacturers and wholesalers are more likely to apply for a new trademark after registering for UPC.

Finally, we merge our UPC adoption measure with industry-level data on U.S. imports and exports (Feenstra, 1996) to study the link between barcodes and trade. Difference-in-difference regressions show a substantial increase in imports within four-digit manufacturing industries where there is more domestic UPC adoption. This result suggests that as retailers adapt to the UPC — by adopting complementary technology, such as scanners, EDI and automated inventory control; carrying a greater variety of products; and even changing their formats — they are more likely to add international suppliers.

This study contributes to several lines of research. First, a number of studies consider the economic impacts of changes in U.S. retailing, starting in the 1980s (e.g., Foster, Haltiwanger, and Krizan, 2006). Within that literature, some authors suggest that bar codes contributed to the emergence of large chains (e.g., Raff and Schmitt, 2016), which in turn stimulated increases in product variety and international trade (Sullivan, 1997; Broda and Weinstein, 2006; Basker and Van, 2010). To our knowledge, this is the first paper to provide evidence directly linking retail technology adoption to trade and product innovation.

Second, there is a small empirical literature on barcodes and scanning. 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. 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.

[2]SKUs are alphanumeric codes that track individual product data at a very granular level within a retail organization. UPCs, which can be used as SKUs, are standardized to allow for inter-firm communication and coordination.

We also contribute to the literature on technology diffusion with network effects, as summarized by Farrell and Klemperer (2007). Specifically, we provide reduced-form evidence of positive externalities between the two sides of the UPC platform: bar codes and scanners. Other studies that measure network effects in two-sided platform adoption include Gandal, Kende, and Rob (2000) for Compact Discs and CD players, and Gowrisankaran, Rysman, and Park (2010) for Digital Video Discs and DVD players.

Because UPC adoption is also a proxy for broader information technology (IT) investments within retail supply chains, this paper fits into a literature that examines how IT affects vertical relationships. Most studies in that literature treat vertical scope as endogenous to IT (Brynjolfsson, Malone, Gurbaxani, and Kambil, 1994; Hitt, 1999; Forman and McElheran, 2012), although recent evidence suggests that intermediate goods often flow across firm boundaries even when vertical relationships exist (Atalay, Hortaçsu, and Syverson, 2014). Within this literature, the closest study to ours is Fort (2017), which shows that IT adoption is associated with supply-chain fragmentation. Here, we assume persistent dis-integration between retailers and their suppliers, and study the impacts of a technology meant to improve arm's-length coordination.

Finally, we contribute to a literature on the diffusion and economic impacts of industry standards. Gross (2017) examines a single large standardization event — the conversion of more 13,000 miles of U.S. railroad track to a standard gauge over one weekend in 1886 — and finds that it led to a sizable redistribution of traffic from steamship to railroad on affected routes. Levinson (2006) provides a history of the shipping container — the standard most often mentioned in conjunction with bar codes. Bernhofen, El-Sahli, and Kneller (2016) find large increases in bilateral trade between countries that have each installed one or more container-ready ports.

The balance of the paper proceeds as follows: Section 2 provides general background on the Universal Product Code and its diffusion. Section 3 describes the data sources and our methods for combining them. In Section 4, we present our analysis of UPC diffusion, including evidence of network effects. Section 5 presents estimates of the impact of UPC adoption on employment, trademarking, and international trade. 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.[3] It was initiated, designed, and implemented by food-industry participants — manufacturers, wholesalers, and retailers — with no government oversight. Unlike the previous major retail innovation of mechanical cash registers, introduced in the 1880s, barcodes 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).[4] 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.[5]

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

[3]Although we use the acronym UPC for simplicity, the Universal Product Code is officially abbreviated "U.P.C." because the certification mark on "UPC" is held by the Uniform Plumbing Code.

[4]The number of digits in the prefix increased from five to six around 1990 (Brown, 1997, p. 191).

[5]UPC adopters were not limited to a single registration, and many large firms registered for multiple prefixes. Because a single prefix could accommodate up to 100,000 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, or the Southern and Western divisions). Thus, the registration cost does not appear to have been prohibitively high.

required manufacturers to redesign their product labels to make room for the symbol, and in some cases to invest in printing technologies that allow for sufficiently precise bars and minimize 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 more slowly than the UPC registration process. Brown (1997, p. 115) reports that by mid-1975, "50 percent (by volume) 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 initially expressed some dismay over the imbalance in adoption rates. For example, the July 1976 edition of *UPC Newsletter* featured an editorial entitled "Time to Get Involved" noting that there had been just 78 retail scanner installations compared to 4,412 manufacturer UPC registrations (Uniform Product Code Council, 1976). One reason for imbalance is that a single UPC registration — sufficient for a firm with 100,000 individual SKUs — was much cheaper than installing scanners at the checkout.[6] 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.[7]

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.[8] Around that time, the major general-merchandise retailers

[6]Basker (2012) estimates the cost of an early scanning system for a multi-lane supermarket at $300,000 in 1982 dollars.

[7]The incremental cost of adding a barcode to a label during its redesign may have been relatively low (Brown, 1997, p. 62).

[8]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. Beck, Grajek, and Wey (2011) study the adoption of scanners in Europe.

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 Walmart's annual reports makes some reference to investments in UPC-based point-of-scale scanning systems. Abernathy and Volpe (2011) report that K-Mart and Walmart required apparel suppliers to place a bar code on every item starting in 1983 and 1987 respectively.

3 Data

To study the diffusion and impact of the UPC, we construct a firm-level panel dataset containing information on UPC registrations, employment, and trademarking. These data contain information on approximately 779,300 manufacturing firms and 866,500 wholesalers over the period 1975 to 1992, comprising 10.7 million firm-year observations. 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 populations, with around 9.6 percent of firms exiting the data set in any given year. Table 1 reports summary statistics for each of the variables described below.

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 number of new registrations per year, and the right panel shows the distribution of a "fee class" variable based on annual revenue (in millions of dollars) of the registering firm

(Zimmerman, 1999, Appendix E).[9] The vast majority of registrants 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 Appendix. Ultimately, we successfully match between 40 and 50 percent of UPC registrations to establishments in three sectors: manufacturing, wholesale, and retail. Figure 2 shows the match rate, by year and by firm size. Our match rate is higher for large firms, and therefore greater in the early years, when new registrants tended to be larger.

There are two main reasons why the overall match is bounded away from 100 percent. First, we only include firms with one or more establishments in the manufacturing, wholesale, and retail sectors. 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.[11] Then, 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.[12]

Table 1 shows that 3.8 percent of the observations in the manufacturers sample and

[9]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. We combine the top three fee classes in the figures to comply with Census rules on disclosure avoidance.

[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 in 1978, after which it stabilizes between 85 and 90 percent.

[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.

[12]For additional details on the aggregation procedure, see the Appendix.

2.2 percent of the observations in the wholesalers sample belong to firms that registered for a UPC. At the firm-year level, the corresponding adoption rates are 1.9 percent and 1.0 percent, suggesting that adopters held a UPC for about half of the years when they appear in the sample. These numbers 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 show in the next section.

Our analysis of network effects uses a measure of UPC adoption by rivals in the same four-digit SIC code as a focal firm.[13] Because firms may have one or more establishments (plants, warehouses, etc.), each with a different SIC, we calculate rival UPC adoption in two steps. The first step computes average adoption within a four-digit SIC code at the establishment level (assuming that all establishments within a given firm adopt at the same time), and the second step aggregates this measure up to the firm level. Specifically, in the first step, we compute the employment-weighted average of competitors' UPC adoption for each establishment e:

$$\widehat{\mathrm{UPC}}_{et} = \sum_{j \in \{s(e) \backslash i(e)\}} \frac{\mathrm{Employment}_{jt}\mathrm{UPC}_{jt}}{\sum_{k \in \{s(e) \backslash i(e)\}} \mathrm{Employment}_{kt}} \tag{1}$$

where $\{s(e) \backslash i(e)\}$ 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.[14] In the second step, we take the employment-weighted average of $\widehat{\mathrm{UPC}}_{et}$ over all establishments in firm i:

$$\overline{\mathrm{UPC}}_{it} = \sum_{k \in e(i)} \frac{\mathrm{Employment}_{kt}\widehat{\mathrm{UPC}}_{kt}}{\mathrm{Employment}_{it}} \tag{2}$$

[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).

[14]In some specifications we define $\{s(e) \backslash i(e)\}$ 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.

Table 1 shows that $\overline{\text{UPC}}_{it}$ averages 16.1 percent across all manufacturing firm-years and 5.1 percent across all wholesaler firm-years in our panel.

We use data on U.S. Trademark (TM) applications to study the link between UPC adoption and product innovation. These data come from the USPTO Trademark Case Files Dataset (Graham, Hancock, Marco, and Myers, 2013), merged to the Business Register via the matching procedure described in Dinlersoz, Goldschlag, Myers, and Zolas (2017).[15] A trademark is a "word, phrase, symbol, design, color, smell, sound, or combination thereof" that identifies products sold by a particular party (15 U.S.C. § 1127). Although TMs need not be registered, federal registration in the U.S. provides *prima facie* evidence of ownership, affords nationwide protection, and is required for enforcement in federal court. Millot (2009) reviews the empirical literature on TMs, and argues that they are a useful indicator of product and marketing innovation.

We use the TM data to construct two outcomes: a count of new TM applications in a given calendar year, and an indicator that a firm applied for at least one new TM. Both variables are based on applications that eventually produce a registered TM.[16] 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.

We analyze scanner diffusion within the grocery sector using data from the Food Marketing Institute, described in Basker (2012). For this analysis, we create a store-specific measure

[15] Matching administrative data on U.S. TM registrations to the Census Business Register is a difficult and time-intensive task, and we are indebted to these authors for making their match available to us.

[16] This restriction is important because, starting in 1989, firms could file "intent to use" applications for trademarks that were never actually used, and we observe a large increase in applications around that time. Registration indicates that the TM was actually used in commerce.

of upstream UPC adoption within a retail supply chain. Because we do not observe actual supply-chain relationships, our measure combines industry-level variation in UPC adoption with store-level data on the merchandise mix offered by individual retailers. Specifically, we create a concordance between upstream industries and broad merchandise lines (e.g. food, women's apparel, furniture), and use this mapping to construct a variable, Upstream_{st}, that is essentially the revenue-weighted average UPC adoption rate of industries supplying store s at time t. For additional details on the creation of this variable, see the Appendix.

Table 2 presents summary statistics for the two store-year samples that we use to analyze scanner diffusion. The left panel contains all food stores (SIC 5411), and the right panel contains only stores that adopt a scanner by 1984.[17] The first row of the table shows that the mean hazard of grocery store scanner adoption between 1974 and 1984 was 0.8 percent. Overall, 3.7 percent of the grocery stores in our sample adopted scanning by the end of the sample period. Our store-specific measure of UPC exposure, Upstream_{st}, averages 0.38 for adopters and 0.36 for all food stores. If we interpret upstream industry-level UPC adoption as the share of barcoded items produced by that industry, the latter number implies that 36% of items in the average food store are barcoded.

Our final data set is an industry-year panel containing measures of UPC adoption and international trade. Specifically, we supplement the 1987 SIC version of the NBER-CES Manufacturing Industry Database with our industry-level adoption measure $\widehat{\text{UPC}}_{jt}$, and merge it with data on U.S. imports and exports by four-digit SIC, based on data collected by Feenstra (1996) and concordances from Pierce and Schott (2009). After combining a small number of industries that have no imports, no exports, or only imports or exports from Canada, with closely related industries (to avoid zero cells when we take logs), this

[17]Because we use these data to estimate hazard models, both panels contain store-year observations up to and including the year of scanner adoption, at which time a store is removed from the sample.

yields a strongly balanced panel of 422 manufacturing industries for the years 1975 to 1992.[18] Documentation and summary statistics for the underlying productivity and trade data sets are available on the NBER web site.[19]

4 Diffusion

4.1 Firm Size and Industry

We start by partitioning all manufacturing firms, in each Economic Census year, by revenue quartile, and calculating the share of firms in each quartile that have registered for a UPC by that year. The registration rates are shown on the left-hand-side panel of Figure 3. Among firms in the largest quartile, approximately two percent registered for a UPC by 1977, and nearly ten percent registered by 1992. Smaller firms have lower registration rates; no more than two 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.[20] Overall registration rates in wholesale are lower, but the ranking of adoption rates by quartile is very similar. Retail adoption rates follow a similar pattern, but at a much lower level, and are omitted from the figure.[21]

[18] Examples of combined industries are SIC 3322 (Malleable iron foundries) with 3325 (Steel foundries, not elsewhere classified), and SIC 3761 (Guided missiles and space vehicles) with 3769 (Space vehicle equipment, not elsewhere classified).

[19] The NBER-CES Manufacturing productivity data are available at `http://www.nber.org/nberces/`. U.S. mports and exports by 1987 SIC are available at `http://faculty.som.yale.edu/peterschott/sub_international.htm`.

[20] Wholesalers may be merchant wholesalers, which are intermediaries that buy inputs and may package, repackage, or label them for sale to retailers, or 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.

[21] Although UPC registration is not necessary to implement scanning, there are several reasons why retailers registered. Some retailers registered to show support for the system. Others owned upstream establishments,

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). Figure 1 shows a bunching of total registrations in 1983, consistent with widespread adoption of the UPC by general-merchandise suppliers around that time.

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.[22] 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 five percent in 1975, and increases to about 20 percent by 1992. The employment share, however, remains fairly stable at 60 percent, reflecting the fact that large firms registered early and later registrants are small.[23] 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) mostly supply their respective retail specialist retailers, as well as 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

---

such as a production facility for private-label items. Grocery stores with meat and deli departments may also register for a UPC in order to print their own barcodes on variable-weight items (Selmeier, 2008, pp. 131–132).

[22]As detailed in the Appendix, firms are classified by their predominant industry. Firms that operate multiple establishments are classified in a single industry despite having some establishments, perhaps a significant number, in other industries.

[23]Food manufacturers include both intermediate- and final-goods manufacturers. We expect the registrations to be disproportionately concentrated among final-goods manufacturers, so these rates may be under-estimates of the registration rates among final-goods producers.

growth because here, too, larger firms adopt earlier.

4.2 Network Effects: UPC

The UPC is a classic case study for two-sided network effects. The basic argument 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 this "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 have only limited information on adoption of complementary technologies, such as barcode scanners, by downstream firms. If we had comprehensive data on supply-chain links and scanner adoption, we could estimate the system of equations

$$\Delta \text{UPC}_{it} = f\left(\mathbb{E}_t\left(\text{Scanner}_{j(i),t+1}\right), \cdot\right) \tag{3}$$

$$\Delta \text{Scanner}_{jt} = g\left(\mathbb{E}_t\left(\text{UPC}_{i(j),t+1}\right), \cdot\right) \tag{4}$$

where the outcomes ΔUPC_{it} and $\Delta \text{Scanner}_{jt}$ are binary variables indicating that manufacturer i registered for a UPC prefix or retailer 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 focal firm's supply chain.[24] The identification problem is that a positive correlation between upstream UPC registrations and downstream scanner adoption can be causal in one or both directions, or may be due to an omitted variable; without instruments we cannot determine the relative importance of the two channels.

Instead of estimating (3) directly, we use a reduced-form approach.[25] When there are indirect or "cross-side" network effects, the benefits of UPC adoption increase with the share of adopting competitors, because the share of adopting competitors is increasing in the share of downstream firms that 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:

$$\text{UPC}_{it} = \lambda_{at} + \beta\overline{\text{UPC}}_{it} + \varepsilon_{it} \tag{5}$$

where the outcome UPC_{it} indicates that firm i registered in year t, and the main explanatory variable $\overline{\text{UPC}}_{it}$ is the employment-weighted rival UPC adoption rate. The model includes a full set of firm-age by calendar-year fixed effects (λ_{at}), and each firm is retained in the data only until the year when it registers, so that β can be interpreted as the change in the

[24] Although in principle, stores could choose to install scanners in some checkout lanes and not others, the choice was binary in practice (Basker, 2012). Brown (1997, p. 116) also describes how optimistic expectations promoted early UPC adoption by grocery manufacturers. Specifically, he 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."

[25] The Appendix shows how to derive our reduced-form specification, up to a monotonic transformation of $\overline{\text{UPC}}$, from a linear system of first-order differential equations analogous to (3) and (4).

hazard of UPC adoption if all of a firm's rivals switched from non-adopters to adopters.[26,27]

Manski (1993) discusses identification for models such as (5), where an individual choice is regressed on a group average for the same variable. We interpret β as what he calls a correlated effect: upstream firms behave similarly because they face a similar institutional environment (specifically, pressure from retailers that have installed scanners). The alternative interpretation, which Manski calls an endogenous effect, is that rival UPC adoption has a direct causal impact on the choice of a focal firm.[28] To reinforce the idea that our estimates capture network effects, we estimate a placebo model that seeks evidence of geographic spillovers. If UPC adoption is mainly caused by downstream scanner adoption (i.e., correlated effects), then we should find much smaller spillovers from geographically proximate firms that sell through different supply chains.

Our estimates of Equation (5) are reported in Table 3. For both manufacturing and wholesale firms, we find statistically significant and economically meaningful spillovers in UPC adoption.[29] These are measured by the coefficients in the first and third column of the table. The standard deviation of the within-industry spillover variable (as reported in Table 1) is 0.17 for manufacturers. Thus, 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. These results could be caused by either a coordinated effort to start the scanner-adoption bandwagon or time-varying industry-level unobservables that increase the returns to adoption, such as a correlated reduction in costs. We find the network effects interpretation more plausible because the costs of UPC adoption were small, and do not

[26] Jenkins (1995) discusses estimation of discrete-time duration models using "panel" data with one observation per-unit per-period until exit (here, UPC registration) and shows that logit models reproduce the likelihood for a proportional hazard specification. We estimate the analogous OLS regression.

[27] 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.

[28] 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.

[29] We have also estimated these regressions using a one-year lag of the RHS variable, $\overline{\mathrm{UPC}}_{i,t-1}$; the estimated coefficient $\hat{\beta}$ falls very slightly.

seem to have a large industry-specific component.

To perform the placebo tests, we re-define $\overline{\text{UPC}}_{it}$ as the average UPC adoption rate by three-digit zipcode, and re-estimate Equation (5). The coefficients in the second and fourth columns of Table 3 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.

The placebo results provide support for the presence of two-sided network effects. An alternative explanation, however, is that firms imitate their direct competitors — but not other nearby firms. In Section 5.1 we show that upstream employment increases following UPC adoption, providing further evidence of vertical interactions (i.e. correlated effects) as opposed to pure imitation.

4.3 Network Effects: Scanners

To measure network effects on the retail side of the UPC platform, we use data on grocery-store scanner adoption from Basker (2012). The analysis is conducted at store (establishment) level rather than chain (firm) level because we have store-level data, and because scanner adoption occurred gradually within large retail chains.[30]

In a simple model of adoption, firms compare the costs of installing scanners in a store to the (expected) benefits of scanning, which depend critically on the share of the store's suppliers that have attached barcodes to their packages. This suggests estimating a version of Equation (4), where the hazard of scanner adoption at store s is a function of upstream UPC adoption:

$$\text{Scanner}_{st} = \lambda_{at} + \beta \text{Upstream}_{st} + \varepsilon_{st} \tag{6}$$

In this model, the coefficient on upstream UPC adoption is identified by two types of

[30]Basker (2015), using data from public sources, calculates that approximately one third of Safeway stores and one half of Kroger stores had installed scanners by the end of 1984.

variation. First, even within grocery retailing, stores differ with respect to the proportions of food, tobacco products, cleaning supplies, and other goods (such as apparel or home furnishings) that they sell. These differences in merchandise mix create cross-sectional variation in Upstream_{st}. Second, holding a store's merchandise mix constant, the gradual diffusion of the UPC creates longitudinal variation in upstream UPC adoption. The store-age by calendar-year fixed effects λ_{at} control for a possible nonlinearity in scanner adoption: because scanner installation typically required a full front-end remodel, the stores most likely to get them were either new establishments or older stores due for a renovation.

We estimate Equation (6) on a sample that includes all stores in SIC 5411 (food stores), and a second sample that only includes stores that adopt scanners by 1984. The second sample allows us to isolate the effects of upstream UPC adoption on the timing of scanner adoption, at the cost of selecting on the outcome variable. For each sample, we first report a minimalist regression in which we control only for store age by calendar year fixed effects, and then a regression that adds measures of employment for both the store and the chain to which it belongs; an indicator for vertical integration (i.e., owning at least one wholesale or manufacturing plant); and an indicator for the owning firm's UPC registration. As in our UPC-adoption regression above, stores that have installed scanners drop out of the sample in subsequent years, so β corresponds to a change in the hazard of adoption. For all models, we cluster standard errors at the store level.

Table 4 presents our estimates. The coefficient $\hat{\beta}$ is positive and statistically significant across all samples and specifications, consistent with the presence of two-sided network effects. To interpret the magnitude of this coefficient, note that a one-standard-deviation increase in Upstream_{st} is associated with a 22 percent increase in the baseline hazard (i.e., the sample mean adoption rate reported at the the bottom of Table 4). The unreported coefficients on firm-level controls indicate increased scanner adoption by larger firms, consistent with the presence of chain-level economies of scale in scanner deployment.

Estimates of $\hat{\beta}$ appear much larger for the sample of scanning stores. However, the mean adoption rate is much higher in this sample (by construction), so a one standard deviation increase in Upstream increases the baseline hazard by only 3.5 percent. Thus, although our measure of upstream UPC adoption is positively correlated with scanner adoption in both samples, it seems to explain more about which stores will install scanners (before 1984) than about how quickly they will do so.

5 Impacts of UPC Adoption

This section estimates the impact of UPC adoption on several outcomes. At the firm level, we analyze employment and product innovation (as measured by new trademarks). At the industry level, we examine the relationship between UPC adoption and international trade.

5.1 Employment

To estimate the impact of UPC adoption on manufacturer and wholesaler employment, we use the following difference-in-difference specification:

$$\ln(\text{Employment}_{it}) = \alpha_i + \lambda_{at} + \beta \text{UPC}_{it} + \varepsilon_{it} \tag{7}$$

where Employment_{it} is firm i's employment in year t, aggregated from establishment-level data in the LBD; α_i is a firm fixed effect; and 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 ε_{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. For each sector, we construct two different estimation samples. The first sample keeps all non-adopting firms as controls, and the second sample matches adopters to non-adopters based on size and employment growth. Thus, in the matched sample analysis, counterfactual employment growth for UPC adopters is estimated by actual employment growth at similarly sized non-adopters that exhibit similar pre-adoption growth rates over the same time period. We interpret matched sample estimates of β as the average treatment effect for treated firms.

The one-to-one matched samples are constructed 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 register for a UPC 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.[31] 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 $(t-1)$.[32] 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.

Although matched firms always share a predominant sector (i.e., manufacturing or wholesale), we do not restrict them to have the same industry for several reasons. First, if UPC adoption is driven by the downstream demand for bar codes, which varies more between industries than within them, matching across industries should reduce concerns about selection on time-varying firm-level unobservables. Second, within-industry non-adopting firms may not be a valid control group because they can be affected by competitors' adop-

[31]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.

[32]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.

tion of the UPC. Third, 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 number and quality of the matches. Finally, time-invariant industry-level heterogeneity is controlled by the inclusion of firm fixed effects.

Coefficient estimates based on Equation (7) are reported in Table 5. 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 estimate for manufacturing implies a 16 percent increase in employment following UPC adoption, whereas the matched sample estimate suggests a 9 percent increase. 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.[33] To get a better handle on this question, we estimate a series of event-study regressions of manufacturing establishment employment on UPC adoption. Our main specification is:

$$\ln(\text{Employment}_{it}) = \alpha_i + \lambda_{at} + \sum_k \beta_k \Delta\text{UPC}_{i,t+k} + \varepsilon_{it} \tag{8}$$

where α_i and λ_{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 β_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

[33] Another issue with our baseline difference-in-difference specification is that survival rates may differ for adopters and non-adopters. If firms that anticipate exit also avoid the costs of UPC adoption, our estimates could be downward biased (assuming that the true outcome for a defunct firm is zero). Unreported hazard model estimates reveal that there is indeed a strong negative correlation between UPC adoption and exit.

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 roughly 10% above the counterfactual level.

The abrupt increase in relative employment observed in these matched difference-in-difference models is a striking result. If UPC registration were primarily a proxy for the adoption of a broader set of technological and organizational changes, it would most likely be associated with a slow-and-steady divergence in employment. The discrete jump suggests to us that manufacturers adopted the UPC specifically 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 demand shocks caused firms to adopt the UPC. Nevertheless, the sudden increase in employment suggests that UPC adoption was a necessary condition for achieving scale through partnering with larger downstream firms.

Though we view employment as a proxy for firm size, we have also examined the link between UPC adoption and revenue.[34] Because for this time period, revenues are available only in five-year intervals from the Economic Census, we cannot estimate an event-study specification, but we do estimate fixed effects models, replacing the LHS variable in Equation (7) 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 those in Table 5, so we do not report them separately.

Network Effects and Employment

We can 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. As in Equation (5), rivals' registrations serve as a proxy for downstream complementary technology adoption, or the expectation thereof.

To implement our test for network effects, we add two variables to the difference-in-difference specification in Equation (7): the employment-weighted UPC adoption rate for each firm's competitors, $\overline{\text{UPC}}_{it}$ (defined as in Equation (2)), and an interaction of $\overline{\text{UPC}}_{it}$ with the focal firm's own adoption indicator. Specifically, we estimate:

$$\ln(\text{Employment}_{it}) = \alpha_i + \lambda_{at} + \beta \text{UPC}_{it} + \gamma \overline{\text{UPC}}_{it} + \delta \text{UPC}_{it} \overline{\text{UPC}}_{it} + \varepsilon_{it} \tag{9}$$

Estimates of the direct effect of adoption, β, and the interaction term, δ, for each of the four samples (OLS and matched, manufacturing and wholesale) are shown in Table 6.

[34] It is easy to see how scanning and EDI can reduce labor costs downstream (see Basker, 2012). It is somewhat harder to tell a story in which the UPC reduces labor demand by producing firms. If there is a labor saving effect, however, then our estimates of increases in employment, which we attribute to increased demand for the firm's output, are a lower bound.

The main effect of UPC adoption (or the impact for the first adopter in an industry, β) 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).[35] 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 by one third — from 6.6 to 8.7 percent — for manufacturers, and from 11.5 to 13.1 percent for wholesalers.[36]

One concern with these estimates is that, as explained earlier, we do not match firms on industry or industry characteristics. As a result, even in the matched sample, the coefficient δ 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 restriction reduces the sample size by 40 percent. Our coefficient estimates change very little for this sample, but estimates of the interaction effect, δ, lose statistical significance due to larger standard errors.

5.2 Trademarks

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

[35] We have also estimated these regressions using a one-year lag of competitors' adoption, $\overline{\text{UPC}}_{i,t-1}$. Results are very similar.

[36] The marginal effect of UPC_{it} is equal to $\beta + \delta\overline{\text{UPC}}_{it}$. For manufacturers we estimate $\delta = 0.123$, and (from Table 1) the standard deviation of $\overline{\text{UPC}}_{it}$ is 0.17, so a one SD shift in $\overline{\text{UPC}}_{it}$ increases the predicted marginal effect by $0.17 \cdot 0.123 = 0.021$. For wholesalers, the predicted change in the marginal effect when rival UPC adoption increases by one SD is $0.073 \cdot 0.215 = 0.016$.

shopping." We investigate this hypothesis using registered TM applications as a proxy for variety-increasing product innovation.

As a starting point, Figure 7 presents aggregate time-series evidence from the grocery sector. The solid line shows annual new product introductions according to the periodical *New Product News*, and the dashed line shows the average number of SKUs per grocery store as reported in *Progressive Grocer*.[37] Both series are "ocular reproductions" of data reported in Sullivan (1997).[38] The dotted line is a count of new TM applications for grocery-related products that we constructed from the USPTO data.[39]

Figure 7 helps motivate a firm-level trademark 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, all three time-series experience a trend break around 1980 — roughly the time period when the UPC was diffusing through the grocery supply chain, as illustrated in Figure 4. This is consistent with the hypothesis that the UPC and related innovations encouraged grocery product proliferation, and begs the question of whether increased trademarking is concentrated among firms that actually registered for a UPC.

Our firm-level TM analysis uses the same two-way fixed effects specification presented in Equation (7), 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 more TMs. The full-sample results are presented in Table 7. Difference-

[37] New products and SKUs per store are not mechanically influenced by scanning. According to Sullivan (1997, p. 474), "Company representatives said that the increase could not be due to changes in their sampling (for example, an increase in the area of the country covered) or to the adoption of scanner systems by supermarkets (neither company relies on scanner-based data sources)."

[38] We resorted to the eyeball method because her original data have been lost. The SKU series has a gap in coverage between 1972 and 1982, as shown in the figure.

[39] In order to restrict our count of TM applications to the grocery industry, we focus on applications with a three-digit Nice code corresponding to food, beverages, pharmaceuticals or paper products. We adjust for missing data in the years before 1977 using a procedure described in the Appendix.

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.[40] In the second column, we observe a six-fold increase in the count of new TM filings after UPC adoption.[41]

As with the employment results, we have estimated event-study specifications and used a matching estimator to address selection. The pattern of results bears strong similarities to our employment findings. Event-study specifications with the full sample show a strong pre-UPC-registration trend in both the probability of a firm having at least one TM registration and in the expected number of TMs. To address the pre-trend, we create a matched sample using a procedure very similar to the one we use for employment.[42] By design, the pre-registration trend disappears in the matched samples. The probability of TM registration increases discretely in the year of UPC adoption, and stays higher. The count of TM registrations follows the same pattern for wholesale firms; however, for the manufacturing sample, there is only a temporary increase in the number of TMs immediately after UPC registration. For difference-indifferences estimates based on Equation (7), matching reduces the estimated treatment effect for three of the four TM outcomes (the exception being wholesalers' probability of having a TM registration, where the estimates increases relative to

[40]This marginal effect is calculated relative to the mean of the outcome, i.e., $\left(\frac{0.053}{0.016} - 1\right)$.

[41]We subjected these results to a variety of robustness checks. First, instead of counting all TM filings that led to a registration, and dating them to the filing year, we create a variable that counts applications for which we have the date of first use in commerce, linking each new TM to that date. Second, we re-estimate each model keeping only the firms that actually applied for at least one TM during our sample period. Finally, we estimate a Poisson specification of (7) to account for the skewed outcome and many zeros. The results of these models are all broadly similar to those reported in Table 7.

[42]Specifically, each firm in the matched sample has a unique match in the same sector (manufacturing or wholesale). For firms observed for the first time in the year of registration, the controls are chosen at random from the set of firms of the same vintage. For firms ages 1–4 at the year of registration, the controls share a vintage and are matched on the cumulative number of TMs they have registered as of year $(t-1)$. For firms aged five and over at the year of registration, the controls are other firms that are at least five years old in year t, by the cumulative number of TMs they have registered between years $(t-5)$ and $(t-1)$. The vast majority of firms have zero cumulative TMs.

the full sample). The matched-sample estimates also remain statistically significant for three of the four outcomes variables, with the exception of the annual TM count for manufacturers.

In summary, our firm-level analysis shows that UPC registration is associated with a sharp economically and statistically significant increase in employment, revenue and trademark registrations for both manufacturers and wholesalers. Interpreting these results requires care. Although we use matching to remove selection effects, UPC adopters may still be more likely to adopt new technologies, use innovative management practices, and grow even in the absence of UPC adoption. Nevertheless, our findings suggest that once downstream technologies are in place, upstream UPC adoption helped manufacturers and wholesalers achieve scale by supplying large retailers. And the trademark results suggest that joint adoption of scanning and bar codes created new opportunities for producing and distributing a wider assortment of goods. The significance of these developments is illustrated by the role that both new retail formats and increased product variety played in later debates over aggregate price and productivity measurement (e.g., Boskin, Dulberger, Gordon, Griliches, and Jorgenson, 1998).

5.3 International Trade

Although several studies examine the link between importing and increased product variety (e.g, Feenstra, 1994; Broda and Weinstein, 2006), there is surprisingly little evidence linking import growth to changes in retail technology or productivity. Nevertheless, several observers such as Basker and Van (2010) and Raff and Schmitt (2016) suggest that technological innovations, including the UPC, were a key factor behind the growth in both imports and modern retail chains. The argument is explained in Basker and Van (2006): technological changes that increase a chain's optimal size and lower its marginal input costs lead to lower prices, which stimulate demand. If contracting with offshore suppliers entails paying a fixed cost to purchase at a lower price, the chain will start importing when it reaches a minimum size threshold, at which point marginal cost again falls, leading to increased profits and pushing the chain to expand still further. Although we cannot test each link in this causal chain, the UPC-registration data allow us to measure an industry-level association between domestic retail technology adoption and international trade.

The outcome variables in our trade analysis are (logged) total U.S. imports and total U.S. exports. Our estimates are based on the following reduced-form difference-in-difference specification:

$$\ln(\text{Trade}_{jt}) = \alpha_j + \lambda_t + \beta\overline{\text{UPC}}_{jt} + \varepsilon_{jt} \tag{10}$$

where Trade_{jt} measures U.S. imports or exports by industry j in year t; $\overline{\text{UPC}}_{jt}$ is employment-weighted industry average domestic UPC adoption; α_j are industry fixed effects; λ_t are calendar-year fixed effects; and the error term ε_{jt} is clustered at the industry level.

Our main explanatory variable, $\overline{\text{UPC}}_{jt}$, is based on adoption by domestic establishments. We do not observe adoption by foreign suppliers. Nevertheless, evidence of network effects in the diffusion process (see Section 4.3) suggests that it is reasonable to interpret domestic UPC registration in upstream industries as a proxy for adoption of scanners and related technology by retailers in the same supply chain.

The coefficient β in Equation (10) measures the association between UPC adoption and trade. For imports, we expect this coefficient to be positive if scanning and supply-chain automation reduce retailers' cost of working with foreign suppliers. However, a positive coefficient could reflect several different mechanisms, including (a) substitution of imported for domestic final goods, (b) an output-expanding effect if retailers can pass through lower prices to consumers, and (c) increased imports of intermediate goods as foreign manufacturers begin to supply components to domestic producers. It is less clear what we should expect for exporting. Because the UPC code is a domestic standard, it is tempting to view exports as a placebo test. In practice, all of our estimates of the relationship between domestic UPC adoption and exports are statistically insignificant and close to zero, so we focus on the import regressions.

Table 8 presents the results of our trade regressions. The first column reports a baseline specification that contains only the UPC adoption variable, along with industry and year fixed effects. The second column adds controls for logged industry value-added, logged capital-labor ratio, and the logged ratio of production to non-production workers in the industry. The third and fourth columns report estimates of the same two specifications, but exclude Canadian imports in order to address concerns that our results might be influenced by the 1988 U.S. Canada Free Trade Agreement. Across all four models, we find a statistically significant positive relationship between domestic UPC adoption and imports.[43] The magnitude of these coefficients implies that a one standard-deviation increase in industry-level UPC adoption is associated with a 6–7 percent increase in imports.[44]

There are several reasons to be cautious about these results. At the micro-level, it is clear that the estimates are not causal: adding numeric labels to domestic firms' packaging should not cause an increase in trade. If we interpret UPC adoption as a proxy for supply-chain automation and the reconfiguration of retail distribution channels, there remains a strong likelihood of selection on the gains to treatment. That is, barcodes were probably adopted first in the industries where they were most useful, such as food, pharmaceuticals and apparel, and would likely have less impact (or perhaps altogether different impacts) when adopted by manufacturers of industrial goods or heavy equipment. Finally, causality could run in either direction. If trade and technology are complementary inputs to the retail production function, an exogenous increase in imports (e.g. due to tariff reductions) should stimulate domestic technology adoption.

In spite of these concerns, we believe that our regressions provide some of the first evidence linking import growth directly to changes in retail technology. Moreover, the cor-

[43] We also tried excluding intermediate goods imports, based on data from Schott (2004) that classifies any HS code containing the word "parts" or a related term as an intermediate (available at `http://faculty.som.yale.edu/peterschott/sub_international.htm`). This led to a small increase in the magnitude of our estimates.

[44] The within-industry standard deviation of $\overline{\mathrm{UPC}}_{jt}$ is approximately 0.125.

relation between imports and domestic UPC adoption also points to the broad impacts of the entire system of technologies supported by the adoption of UPCs and scanning.

6 Concluding Remarks

Barcodes 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 big-box format (e.g., Holmes, 2001, and Dunlop, 2001). This paper is the first to measure the effects of UPC adoption on upstream employment, product innovation, and industry-level imports, providing a natural complement to the literature on retail productivity (e.g., Foster, Haltiwanger, and Krizan, 2006) and a new addition to the empirical literature on the effects of adopting industry standards.

We show that early 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 in other sectors. We provide new evidence on this point by estimating a reduced-from model of within-industry spillovers in UPC adoption, and find positive effects for both manufacturers and wholesalers. We also find that grocery retailers whose merchandise mix exhibited greater upstream UPC adoption were more likely to install scanners in the early years.

Our investigation of the impacts of UPC adoption suggests that both upstream and downstream firms benefited on several margins. For manufacturers and wholesalers, we find that both revenue and employment increased concurrent with adoption, and stayed higher, consistent with receiving larger orders from retailers. 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 being simply a proxy for adoption of a suite of UPC-related technologies and business practices, which would yield more gradual and consistent increases in employment.

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 within that distribution channel. Moreover, 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. Finally, the positive correlation between UPC adoption and industry-level imports points to broader effects of the entire barcode system, including its role in enabling automated inventory tracking and replenishment, which encouraged large retail chains to seek out more international suppliers.

References

Abernathy, F. H., J. T. Dunlop, J. H. Hammond, and D. Weil (1995) "The Information-Integrated Channel: A Study of the U.S. Apparel Industry in Transition," *Brookings Papers on Economic Activity, Microeconomics*, pp. 175–246.

——— (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.

Abernathy, F. H., and A. P. Volpe (2011) "Technology and Public Policy: The Preconditions for the Retail Revolution," in *The Market Makers: How Retailers are Reshaping the Global Economy*, ed. by G. G. Hamilton, M. Petrovic, and B. Sauer, chap. 2, pp. 50–79. Oxford University Press.

Atalay, E., A. Hortaçsu, and C. Syverson (2014) "Vertical Integration and Input Flows," *American Economic Review*, 104(4), 1120–1148.

Basker, E. (2012) "Raising the Barcode Scanner: Technology and Productivity in the Retail Sector," *American Economic Journal: Applied Economics*, 4(3), 1–29.

——— (2015) "Change at the Checkout: Tracing the Impact of a Process Innovation," *Journal of Industrial Economics*, 63(2), 339–370.

Basker, E., S. Klimek, and P. H. Van (2012) "Supersize It: The Growth of Retail Chains and the Rise of the "Big Box" Store," *Journal of Economics and Management Strategy*, 21(3), 541–582.

Basker, E., and P. H. Van (2006) "Putting a Smiley Face on the Dragon: Wal-Mart as Catalyst to U.S.-China Trade," unpublished paper, University of Missouri.

——— (2010) "Imports 'Я' Us: Retail Chains as Platforms for Developing-Country Imports," *American Economic Review Papers and Proceedings*, 100(2), 414–418.

Bass, F. M. (1969) "A New Product Growth Model for Consumer Durables," *Management Science*, 15(5), 215–227.

Beck, J., M. Grajek, and C. Wey (2011) "Estimating Level Effects in Diffusion of a New Technology: Barcode Scanning at the Checkout Counter," *Applied Economics*, 43(14), 1737–1748.

Bernhofen, D., Z. El-Sahli, and R. Kneller (2016) "Estimating the Effects of the Container Revolution on World Trade," *Journal of International Economics*, 98(1), 36–50.

Boskin, M. J., E. R. Dulberger, R. J. Gordon, Z. Griliches, and D. W. Jorgenson (1998) "Consumer Prices, the Consumer Price Index, and the Cost of Living," *Journal of Economic Perspectives*, 12(1), 3–26.

Broda, C., and D. E. Weinstein (2006) "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121(2), 541–585.

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

Brynjolfsson, E., T. Malone, V. Gurbaxani, and A. Kambil (1994) "Does Information Technology Lead to Smaller Firms?," *Management Science*, 40(12), 1628–1644.

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.

Dinlersoz, E. M., N. Goldschlag, A. Myers, and N. Zolas (2017) "An Anatomy of Trademarking by Firms in the United States," unpublished paper, U.S. Census Bureau.

Distribution Codes, Inc. (1974) "Directory of UGPCC Members," Membership as of July 31, 1974.

Dunlop, J. T. (2001) "The Diffusion of UCC Standards," in *Twenty-Five Years behind Bars*, ed. by A. L. Haberman, chap. 2, pp. 12–24. Harvard University Press, Cambridge, MA.

Dunlop, J. T., and J. W. Rivkin (1997) "Introduction," in Stephen A. Brown, *Revolution at the Checkout Counter: The Explosion of the Bar Code*, pp. 1–38. Harvard University Press, Cambridge, MA.

Farrell, J., and P. Klemperer (2007) "Coordination and Lock-in: Competition with Switching Costs and Network Effects," in *Handbook of Industrial Organization*, ed. by M. Armstrong, and R. Porter, vol. 3, chap. 31, pp. 1970–2056. Elsevier B.V.

Feenstra, R. C. (1994) "New Product Varieties and the Measurement of International Prices," *American Economic Review*, 84(1), 157–177.

——— (1996) "NBER Trade Database, Disk1: U.S. Imports, 1972-1994: Data and Concordances," NBER Working Paper 5515.

Forman, C., and K. McElheran (2012) "Information Technology and Boundary of the Firm: Evidence from Plant-Level Data," Harvard Business School Working Paper 12-092.

Fort, T. C. (2017) "Technology and Production Fragmentation: Domestic versus Foreign Sourcing," *Review of Economic Studies*, 84(2), 650–687.

Fort, T. C., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2013) "How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size," *IMF Economic Review*, 61(3), 520–559.

Foster, L., J. Haltiwanger, and C. J. Krizan (2002) "The Link between Aggregate and Micro Productivity Growth: Evidence from Retail Trade," NBER 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.

Ganapati, S. (2016) "The Modern Wholesaler: Global Sourcing, Domestic Distribution, and Scale Economies," unpublished paper, Yale University.

Gandal, N., M. Kende, and R. Rob (2000) "The Dynamics of Technological Adoption in Hardware/Software Systems: The Case of Compact Disc Players," *RAND Journal of Economics*, 31(1), 43–61.

Gowrisankaran, G., M. Rysman, and M. Park (2010) "Measuring Network Effects in a Dynamic Environment," NET Institute Working Paper 10-03.

Graham, S., G. Hancock, A. Marco, and A. F. Myers (2013) "The USPTO Trademark Case Files Dataset: Descriptions, Lessons, and Insights," United States Patent and Trademark Office.

Gross, D. (2017) "The Ties that Bind: Railroad Gauge Standards, Collusion, and Internal Trade in the 19th Century U.S.," Harvard Business School Working Paper 17-044.

Haltiwanger, J., R. Jarmin, and J. Miranda (2013) "Who Creates Jobs? Small vs. Large vs. Young," *Review of Economics and Statistics*, 95(2), 347–361.

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

Hitt, L. (1999) "Information Technology and Firm Boundaries: Evidence from Panel Data," *Information Systems Research*, 10(2), 134–149.

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

Hwang, M., and D. Weil (1998) "The Diffusion of Modern Manufacturing Practices: Evidence from the Retail-Apparel Sectors," U.S. Census Bureau Center for Economic Studies Working Paper 97-11.

Jarmin, R. S., and J. Miranda (2002) "The Longitudinal Business Database," U.S. Census Bureau Center for Economic Studies Working Paper 02-17.

Jenkins, S. P. (1995) "Easy Estimation Methods for Discrete-time Duration Models," *Oxford Bulletin of Economics and Statistics*, 57(1), 129–138.

Kmart (1982) *Kmart Corporation 1981 Annual Report*. Troy, MI.

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.

Levinson, M. (2006) *The Box: How the Shipping Container Made the World Smaller and the World Economy Bigger*. Princeton University Press, Princeton, NJ.

Mann, M. L. (2001) "Cracking the Code," in *Twenty-Five Years behind Bars*, ed. by A. L. Haberman, chap. 8, pp. 93–106. Harvard University Press, Cambridge, MA.

Manski, C. F. (1993) "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60(3), 531–542.

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

Millot, V. (2009) "Trademarks as an Indicator of Product and Marketing Innovations," OECD Science, Technology and Industry Working Papers, 2009/06.

Mobius, M., and R. Schoenle (2006) "The Evolution of Work," NBER Working Paper 12694.

Pierce, J., and P. Schott (2009) "A Concordance Between Ten-Digit U.S. Harmonized System Codes and SIC/NAICS Product Classes and Industries," NBER Working Paper 15548.

Raff, H., and N. Schmitt (2016) "Manufacturers and Retailers in the Global Economy," *Canadian Journal of Economics*, 49(2), 685–706.

Schott, P. (2004) "Across-Product versus Within-Product Specialization in International Trade," *Quarterly Journal of Economics*, 119(2), 647–678.

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

Sieling, M., B. Friedman, and M. Dumas (2001) "Labor Productivity in the Retail Trade Industry, 1987-99," *Monthly Labor Review*, December, 3–14.

Sullivan, M. W. (1997) "Slotting Allowances and the Market for New Products," *Journal of Law and Economics*, 40(2), 461–494.

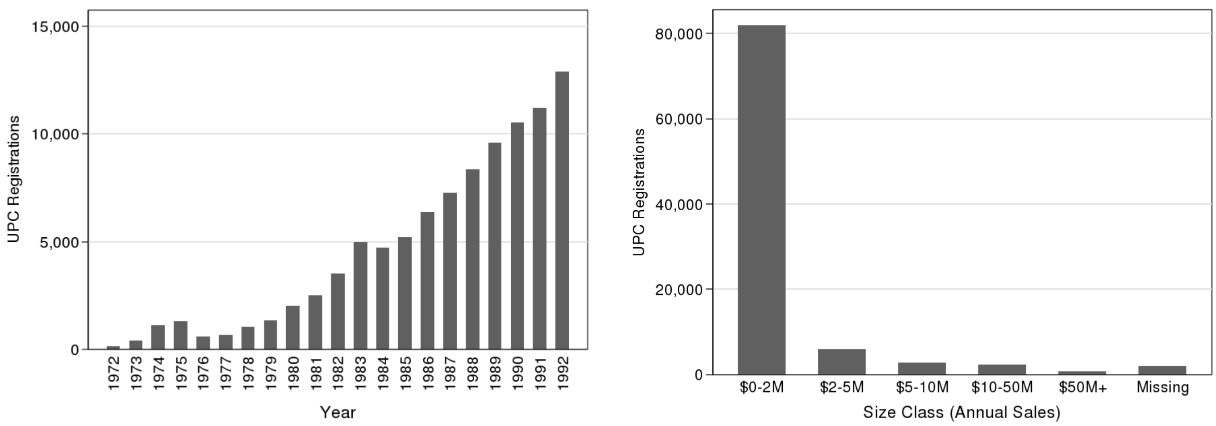
Uniform Product Code Council (1976) "Time to Become Involved," *UPC Newsletter*, 8-6, August 1976, 1–3.

U.S. Census Bureau (2005) *2002 Economic Census: Wholesale Trade: Geographic Area Series*.

Wilson, Jr., T. W. (2001) "How a Low-Tech Industry Pulled off the U.P.C. Standard," in *Twenty-Five Years behind Bars*, ed. by A. L. Haberman, chap. 1, pp. 1–11. Harvard University Press, Cambridge, MA.

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.

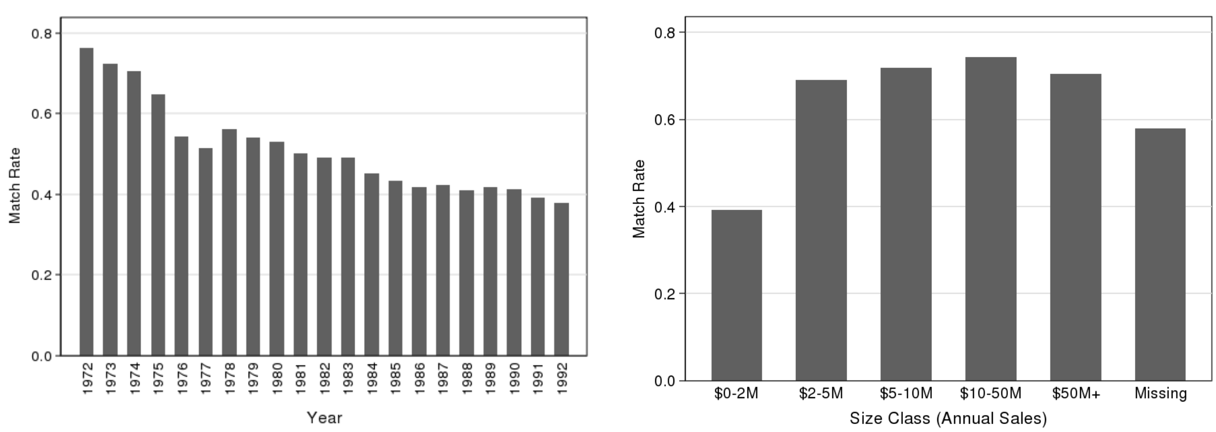
Figure 1. New UPC Registrations



(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

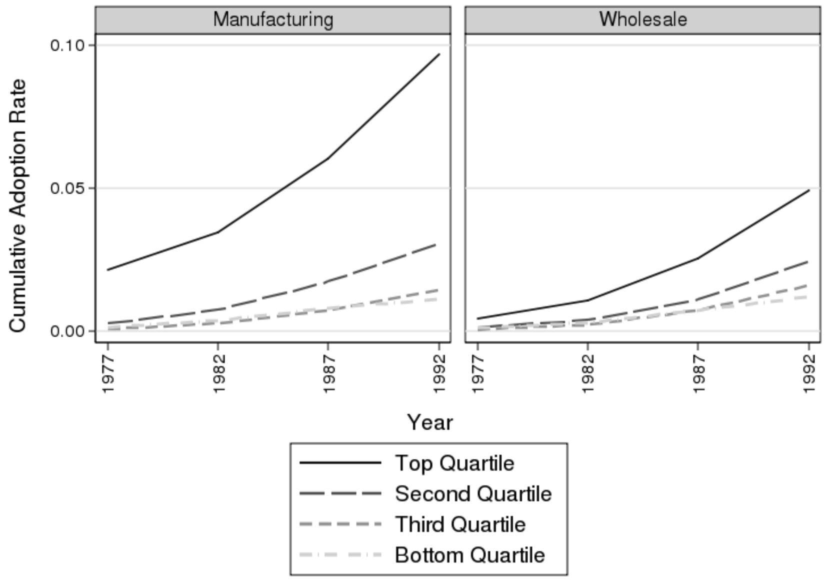


Figure 4. UPC Diffusion for Selected Two-Digit SIC Manufacturing Industries

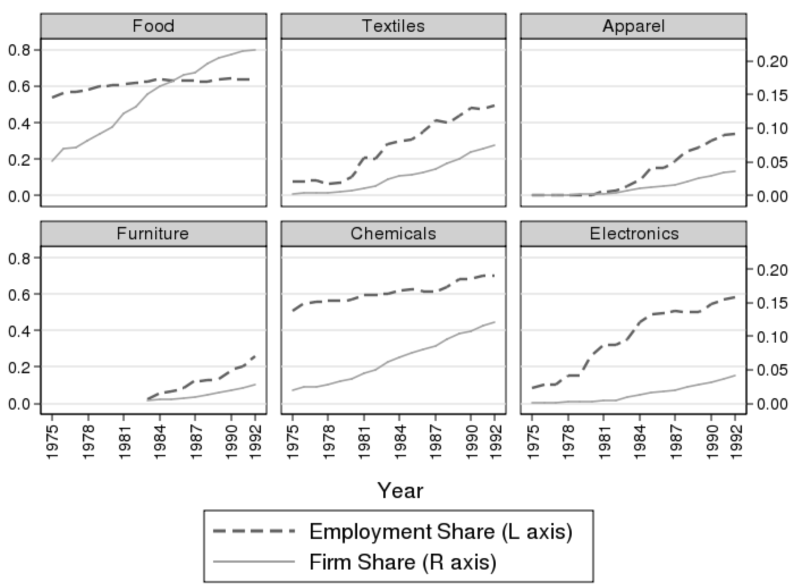
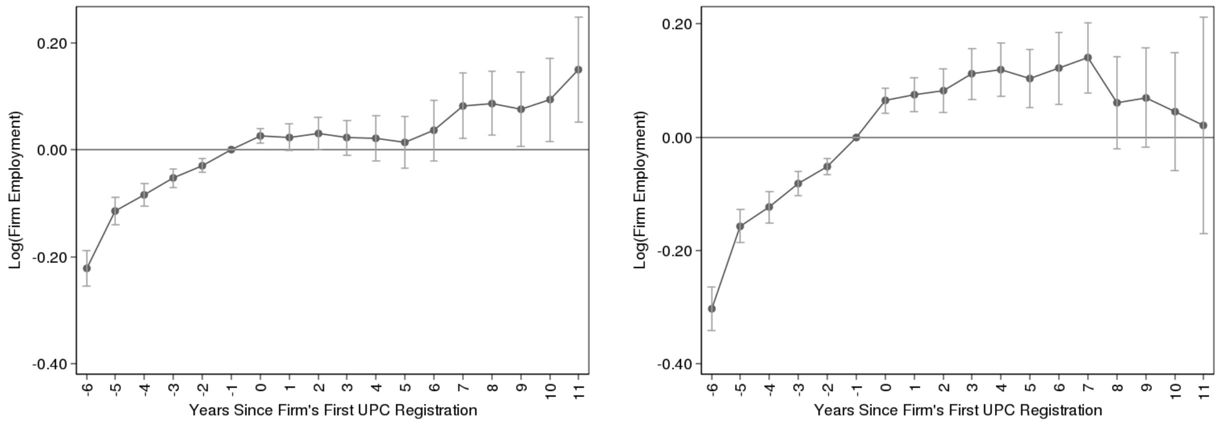


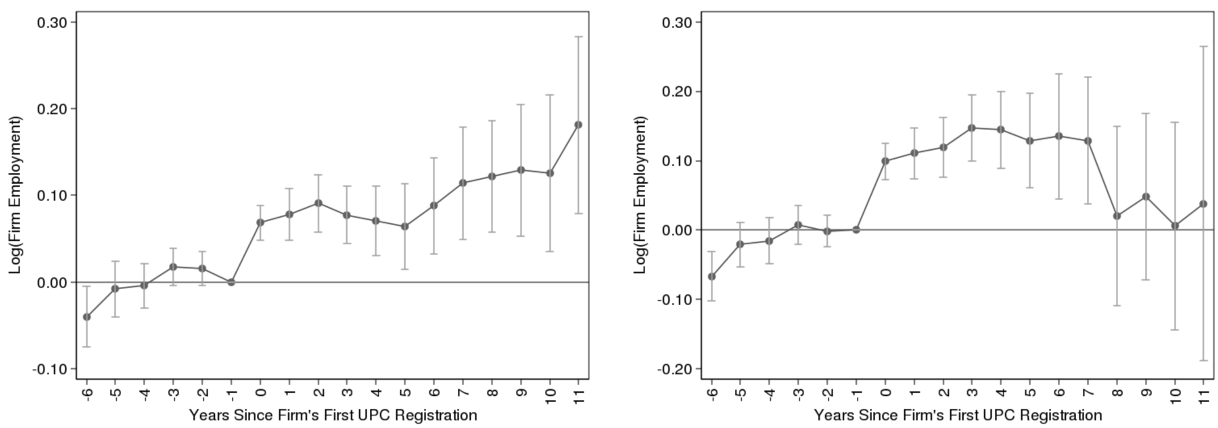
Figure 5. Event Study Coefficients: Manufacturing and Wholesale Employment



(a) Manufacturing (b) Wholesale

Note: Vertical bars represent 95 percent confidence intervals

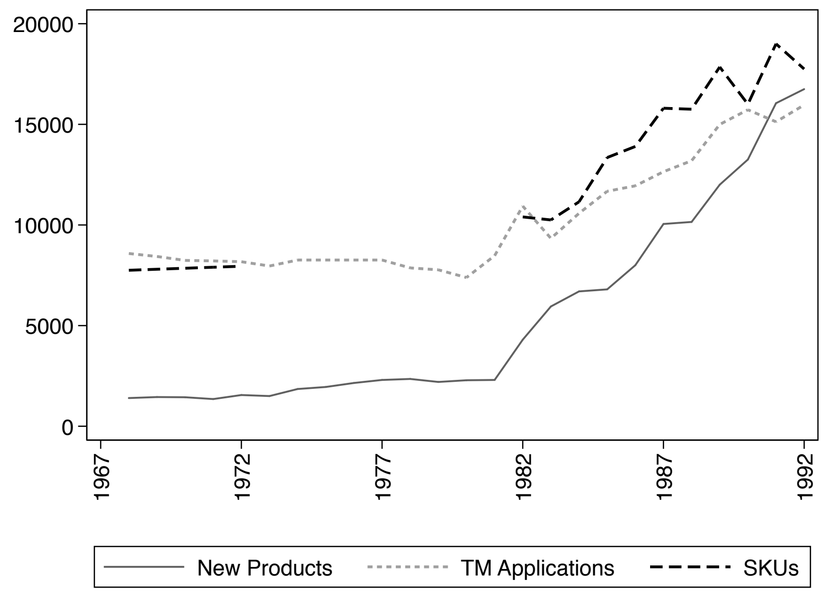
Figure 6. Event Study Coefficients: Manufacturing and Wholesale Employment, Matched Samples



(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, Manufacturing and Wholesale

	Manufacturing		Wholesale	
	Mean	SD	Mean	SD
Employees	72.8	1,540	14.9	193
$\mathbb{I}$[EverUPC]	0.038	0.190	0.022	0.145
$\mathbb{I}$[UPC]	0.019	0.135	0.010	0.100
$\overline{\text{UPC}}_{it}$[a]	0.161	0.170	0.051	0.073
$\mathbb{I}$[EverTM]	0.081	0.266	0.041	0.195
$\mathbb{I}$[TM]	0.016	0.117	0.007	0.081
TM count	0.035	0.614	0.012	0.212
$\mathbb{I}$[Exit|alive at t-1]	0.097	0.296	0.096	0.295
Firms[b]	779,300		866,500	
Observations[b]	5,112,400		5,621,800	

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. Firm and observation counts rounded to the nearest hundred to comply with Census rules on disclosure avoidance.

Table 2. Summary Statistics, Retail

	Food Stores		Scanning Stores	
	Mean	SD	Mean	SD
Scanner	0.008	0.088	0.213	0.409
$\mathbb{I}$[EverScan][a]	0.037	0.188	1.000	0.000
Upstream	0.356	0.071	0.376	0.027
Stores[b]	92,300		3,500	
Observations[b]	437,300		16,000	

Notes: Scanning stores installed a front-end scanner by 1984 (Basker, 2012). Food stores include all scanning stores and other stores in SIC 5411.
[a]EverScan is equal to 1 if the store installed a scanner by 1984.
[b]An observation is a store-year. Firm and observation counts rounded to the nearest hundred to comply with Census rules on disclosure avoidance.

Table 3. UPC Adoption: Industry vs. Geographic Spillovers

	Manufacturing		Wholesale	
Spillover	SIC	ZIP	SIC	ZIP
$\overline{\text{UPC}}_{it}$	0.0282	0.0005	0.0277	0.0050
	(0.0022)***	(0.0003)	(0.0062)***	(0.0003)***
Mean adoption rate	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.
Robust SEs clustered by four-digit firm SIC or three-digit firm ZIP in parentheses.
* p<10%; ** p<5%; *** p<1%
All regressions include firm age×year fixed effects.
[a]Observation counts rounded to the nearest hundred to comply with Census rules on disclosure avoidance.

Table 4. Scanner Adoption

	Food Stores		Scanning Stores	
Upstream_{st}	0.0612***	0.0243***	0.3106**	0.2723**
	(0.0014)	(0.0012)	(0.1308)	(0.1337)
Controls[a]		✓		✓
Mean adoption rate	0.0078		0.2126	
Observations[b]	437,000		16,300	

Notes: Outcome: Scanner adoption. Scanning stores adopt scanners by 1984. Stores remain in sample until year of scanner adoption. Only adoption by food stores is identified in the data.
Robust SEs clustered by store in parentheses. * p<10%; ** p<5%; *** p<1%
All regressions include store age×year fixed effects.
[a]Controls include log store employment, log firm employment, vertical-integration indicator, and own-UPC registration indicator.
[b]Observation counts rounded to the nearest hundred to comply with Census rules on disclosure avoidance.

Table 5. Difference-in-Difference Regressions: Log Employment

	Manufacturing		Wholesale	
	Full	Matched	Full	Matched
UPC_{it}	0.161	0.094	0.228	0.135
	(0.012)***	(0.012)***	(0.014)***	(0.010)***
Observations[a]	5,112,400	221,600	5,621,800	166,500

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

Table 6. Heterogeneity by Industry UPC Intensity: Log Employment

	Manufacturing		Wholesale	
	Full	Matched	Full	Matched
UPC_{it}	0.127	0.046	0.190	0.104
	(0.023)***	(0.024)*	(0.018)***	(0.016)***
$\overline{\text{UPC}}_{it} \cdot \text{UPC}_{it}$	0.091	0.123	0.244	0.215
	(0.056)	(0.058)**	(0.107)**	(0.085)**
Observations[a]	5,112,400	221,600	5,621,800	166,500

Notes: Robust SEs clustered by four-digit firm SIC in parentheses.
* p<10%; ** p<5%; *** p<1%.
$\overline{\text{UPC}}_{it}$ is the adoption rate at the firm's competitors, by SIC
All regressions include $\overline{\text{UPC}}_{it}$, firm, and firm age×year fixed effects.
[a]Observation counts rounded to the nearest hundred to comply with Census rules on disclosure avoidance.

Table 7. Difference-in-Difference Regressions: Trademarking

	Manufacturing		Wholesale	
	Any TM	TM Count	Any TM	TM Count
UPC_{it}	0.053	0.211	0.036	0.079
	(0.004)***	(0.025)***	(0.003)***	(0.010)***
Mean Outcome	0.016	0.035	0.007	0.012
Observations[a]	5,112,400		5,621,800	

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

Table 8. Industry-Level U.S. Imports

	From All Countries[a]		Excluding Canada	
$\overline{\text{UPC}}_{jt}$	0.5389***	0.5204***	0.6521***	0.6345***
	(0.1373)	(0.1346)	(0.1514)	(0.1509)
Controls[a]		✓		✓
Observations	7,596	7,596	7,596	7,596

Notes: Outcome: log(U.S. imports). Robust SEs clustered by four-digit SIC in parentheses. * p<10%; ** p<5%; *** p<1%.
All regressions include industry and year fixed effects.
[a]Controls include log (domestic) industry value added, log capital-labor ratio, and log production-to-non-production worker ratio.

Appendix

Matching UPC Registration to Census Establishment

Our matching procedure begins with a universe of establishments available to match. These establishments come from either the Economic Census (1977, 1982, 1987, and 1992) or the Longitudinal Business Database (1975 to 1992).

The Economic Census is quinquennial (every five years) and covers, with few exceptions, all business establishments with paid employees in the United States.[45] 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.[46] 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.[47] 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

[45]Exceptions include most government-owned or -operated establishments, establishments operated by religious organizations, and agricultural establishments.

[46]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.

[47]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 have different Standard Industrial Classification (SIC) codes.[48] To aggregate the SIC to the

[48] 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.

firm level, we use the firm's employment, payroll (deflated by the CPI), and establishment count 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. In Economic Census years 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.

Upstream UPC Adoption for Stores

Each store in the Census of Retail Trade (CRT) is asked to report its revenue by broad merchandise line (examples of broad lines are food, women's apparel, and furniture). We create a concordance specifying, for each broad merchandise line, the upstream (manufacturing and wholesale) four-digit SIC codes supplying that line.[49] To create a store-level measure of upstream UPC adoption, we start with the industry-level adoption rate $\widehat{\text{UPC}}_{jt}$ for each four-digit SIC j, constructed as in Equation (1), except that we no longer need to omit a focal firm from the calculation. Using our many-to-one correspondence between SIC codes and broad merchandise lines, we then compute an adoption rate $\widetilde{\text{UPC}}_{mt}$ for each merchandise line m. The explicit formula in Economic Census years is $\widetilde{\text{UPC}}_{mt} = \frac{\sum_{k\in j(m)} \widehat{\text{UPC}}_{kt}\text{Employment}_{kt}}{\sum_{k\in j(m)} \text{Employment}_{kt}}$, where $j(m)$ is the set of four-digit industries supplying merchandise-line m, and Employment_{kt} is total employment at all establishments with SIC k in year t.

Finally, for each store s, we compute a revenue-weighted UPC adoption rate across all lines $m(s)$ sold by that establishment:[50]

$$\text{Upstream}_{st} = \sum_{k\in m(s)} \frac{\text{Revenue}_{kt}\widetilde{\text{UPC}}_{kt}}{\text{Revenue}_{st}} \tag{A-1}$$

[49]The merchandise lines are described in more detail in Basker, Klimek, and Van (2012). Our concordance is available online at `http://people.bu.edu/tsimcoe/data/`.

[50]In intercensal years, we use the stores' revenue shares from the prior CRT.

Aggregate TM Registrations in the Grocery Sector

Figure 7 plots the total number of registered U.S. trademark applications in grocery-related categories between 1968 and 1992. The data come from the USPTO Trademark Case Files dataset. However, to create the figure we made several adjustments in order to limit our counts to grocery categories and to account for missing data prior to 1977.[51]

Every TM application is assigned one or more primary classifications that indicates a field of use. In 1973, the U.S. trademark classification system was replaced by the international "Nice Codes" that specify 45 possible primary classes, and all existing U.S. TMs were assigned one or more international codes. In order to restrict our count of TM applications to the grocery industry, we focus on applications with a three-digit Nice code corresponding to food, beverages, pharmaceuticals or paper products. For applications that match to more than one Nice code, we randomly select a single "primary" code for the analysis.

The USPTO Trademark Case Files data show a sharp increase in applications around 1977, because TMs abandoned prior to 1977 were not recorded in the USPTO computer systems (Graham, Hancock, Marco, and Myers, 2013, p. 32). As a result, the raw trend in total 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 our time-series of grocery-related TM registrations. First, we inflate values for 1973 and prior years by multiplying the actual TM count by the ratio of 1977 to 1974 TMs, on the assumption that the share of abandoned TMs remains constant. To be precise, if TM_t^g denotes grocery trademarks applications in year t, we create a new variable:

$$\widehat{\mathrm{TM}}_t^g = \mathrm{TM}_t^g \times \frac{\mathrm{TM}_{1977}^g}{\mathrm{TM}_{1974}^g}$$

and replace TM_t^g with $\widehat{\mathrm{TM}}_t^g$ for all years prior to 1975. Second, we linearly interpolate values

[51]Data and code for replicating Figure 7 are available at `http://people.bu.edu/tsimcoe/data.html`.

for 1975 and 1976 – the filing years most influenced by the change in USPTO record keeping – and plot the resulting time-series in Figure 7.

Neither of these adjustments is made to the data used in our firm-level TM analysis, which includes both grocery- and non-grocery manufacturers and wholesalers, and uses calendar-year fixed effects to control for the change in PTO procedures.

Derivation of Diffusion Models (5) and (6)

Our evidence of network effects in UPC and Scanner adoption is based on reduced form estimates, using a specification derived from Equations (3) and (4). For ease of exposition, we analyze a continuous-time version of this model.

Let M denote the total number of manufacturers and m_t the number of manufacturers that have registered for a UPC by time t. Similarly, we let S be the population of stores and s_t the number that have installed scanners. We assume the functions $f(\cdot)$ and $g(\cdot)$ are linear, and that agents are myopic, so they decide whether to register for a UPC (or adopt scanning) based on the current installed base of complements, without considering future adoption. This leads to the following system of first-order differential equations:

$$\dot{m}_t = \alpha_m s_t \tag{A-2}$$
$$\dot{s}_t = \alpha_s m_t \tag{A-3}$$

Given initial conditions $m_0 = s_0 = 1$, and focusing on the solution where both UPC and scanner adoption increase with time, we have

$$m_t = \exp\{(\alpha_m \alpha_s)^{\frac{1}{2}} t\} \tag{A-4}$$
$$s_t = \left(\frac{\alpha_s}{\alpha_m}\right)^{\frac{1}{2}} \exp\{(\alpha_m \alpha_s)^{\frac{1}{2}} t\} \tag{A-5}$$

Substituting Equations (A-4) and (A-5) into Equation (A-2) yields the following reduced-form specification for UPC adoption in levels:

$$\dot{m}_t = (\alpha_m \alpha_s)^{\frac{1}{2}} m_t \tag{A-6}$$

Finally, to find the hazard rate, we divide both sides of this equation by $(M - m_t)$, the number of manufacturers that have not yet registered for a UPC, yielding:

$$h(m_t) = \frac{\dot{m}_t}{M - m_t} = (\alpha_m \alpha_s)^{\frac{1}{2}} \frac{m_t}{M - m_t} = (\alpha_m \alpha_s)^{\frac{1}{2}} \frac{\overline{\text{UPC}}}{1 - \overline{\text{UPC}}} \tag{A-7}$$

where the last equality relies on the fact that $\overline{\text{UPC}} = \frac{m_t}{M}$.

Equation (A-7) is nearly identical to the reduced-form specification in Equation (5), except that we replace the nonlinear function $\frac{\overline{\text{UPC}}}{1-\overline{\text{UPC}}}$ with $\overline{\text{UPC}}$. This makes our specification equivalent to the well-known Bass (1969) diffusion model, and more importantly, simplifies the interpretation of the coefficient on $\overline{\text{UPC}}$.

We have estimated all of the regressions in Table 3 using $\frac{\overline{\text{UPC}}}{1-\overline{\text{UPC}}}$ as the explanatory variable. This naturally leads to smaller coefficients, but does not change the overall pattern of results, i.e., positive and significant effects when $\overline{\text{UPC}}$ is calculated within industry, and much smaller and/or statistically insignificant coefficients when $\overline{\text{UPC}}$ is calculated within three-digit ZIP codes.

Deriving Equation (6) is much simpler because we have a direct measure of complementary UPC registrations: Upstream_{st}. Dividing both sides of (A-3) by $(S - s_t)$ yields

$$h(s_t) = \frac{\dot{s}_t}{S - s_t} = \alpha_s \cdot \frac{m_t}{S - s_t} \approx \frac{\alpha_s}{S} \cdot \text{Upstream}_t \tag{A-8}$$

where the last step relies on the fact that scanner adoption was negligible, so $s_t \approx 0$, even among food stores. Specifically, Table 2 indicates that just 3.7 percent of the observations in our food-store panel are from stores that installed scanners by 1984.