

NBER WORKING PAPER SERIES

UPSTREAM, DOWNSTREAM:
DIFFUSION AND IMPACTS OF THE UNIVERSAL PRODUCT CODE

Emek Basker
Timothy Simcoe

Working Paper 24040
<http://www.nber.org/papers/w24040>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2017, Revised July 2019

Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau or the National Bureau of Economic Research. All results have been reviewed to ensure that no confidential information is disclosed. The Disclosure Review Board release number is DRB-B0057-CED-20190530. 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, Nathan Chan, 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, Marc Rysman, Martha Stinson, Mary Sullivan, Dan Trefler, Kirk White, Zoltan Wolf, and seminar participants at the U.S. Census Bureau, MINES-ParisTech, Seoul National University, Cornell, Queen's University, Bocconi University, University of Massachusetts, Harvard Business School, Toulouse School of Economics, KU Leuven, LMU Munich, 2017 AEA (Chicago), 2017 IIOC (Boston), 2017 FSRDC Conference (Los Angeles), 2017 CAED Conference (Seoul), and 2018 Israeli IO Day (Jerusalem) for helpful comments and conversations.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2017 by Emek Basker and Timothy Simcoe. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Upstream, Downstream: Diffusion and Impacts of the Universal Product Code
Emek Basker and Timothy Simcoe
NBER Working Paper No. 24040
November 2017, Revised July 2019
JEL No. L11,L15,L81,O33

ABSTRACT

We match 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 firm size and trademark registrations increase following UPC adoption by manufacturers. Industry-level import penetration also increases 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.

Emek Basker
US Census Bureau
Center for Economic Studies
4600 Silver Hill Road
Washington, DC 20233-9100
emek@emekbasker.org

Timothy Simcoe
Boston University
Questrom School of Business
595 Commonwealth Avenue
Boston, MA 02215
and NBER
tsimcoe@bu.edu

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).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 trademarking, 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 producers supplying “UPC-ready” retailers, and vice versa. Retailers become UPC-ready by installing scanners, and by developing electronic data interchange (EDI) capabilities, including electronic payments (Abernathy, Dunlop, Hammond, and Weil, 1999; Basker, 2012).

¹Tim Harford even named the barcode one of “50 Things That Shaped the Modern Economy.” The other economy-shaping discoveries, inventions, and innovations include paper, the bicycle, and antibiotics. For his explanation of this choice, see <http://www.bbc.co.uk/programmes/p04k0066>.

Although we lack comprehensive data on scanner adoption, we provide indirect evidence of network effects by showing that there are positive spillovers in upstream UPC adoption. Specifically, manufacturers become more likely to register for a UPC when other manufacturing firms that use the same retail distribution channels also register.

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 from the manufacturing sector show that UPC registration is associated with increased revenue and employment. We discuss several possible mechanisms for this result, including that firms select into UPC registration due to anticipated demand shocks; that UPC registration coincides with 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. An event-study specification shows that manufacturer employment increases by ten percent in the year of initial UPC registration, and another ten percent over the next two years. We argue that this pattern is consistent with a combination of selection on positive demand shocks and business diversion by manufacturers that can integrate into large retailers' supply chains more readily after registering for a UPC. We also show that the increases in revenue and employment following UPC registration are greater when there is more UPC adoption by other firms selling through the same retail channels, consistent with the presence of network effects.

Aggregate 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 dramatically in the early 1980s, just as scan-

ners were becoming widespread.² To study the relationship between UPC registration and product innovation, we exploit a new link between Census records and the U.S. Patent and Trademark Office (USPTO) Trademark Case Files Dataset (Dinlersoz, Goldschlag, Myers, and Zolas, forthcoming). Difference-in-difference and event study regressions show that individual manufacturers become more likely to apply for a new trademark after registering for a 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 import penetration within four-digit manufacturing industries where there is more domestic UPC adoption. This finding 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 store formats — they are more likely to add international suppliers.

Our investigation of the diffusion and impacts of the UPC 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 barcodes contributed to the emergence of large chains (e.g., Raff and Schmitt, 2016; Basker, Klimek, and Van, 2012), 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 product innovation and trade.

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

²SKUs 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.

beverage companies 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, and varied considerably by industry within the manufacturing sector.

Third, we 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: barcodes 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.

Fourth, because UPC adoption is a proxy for broader information technology (IT) investments within retail supply chains, this study fits into a literature on IT and vertical relationships. Many studies in this 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 within vertically integrated firms (Atalay, Hortaçsu, and Syverson, 2014). Fort (2017) has recently shown that IT adoption is associated with supply-chain fragmentation. Rather than asking how UPC adoption influenced the organization of firms or their supply chains, this paper considers how supply chains influenced UPC diffusion, and also measures the broader impacts of UPC adoption on product variety and trade.

Finally, we contribute to a literature on the diffusion and economic impacts of industry standards. There are many descriptive accounts of standardization. For example, Levinson (2006) provides a history of the shipping container, and the volume edited by Greenstein and Stango (2007) contains several other case studies. Quantitative studies on the causal impacts of standards adoption are less common. Two recent exceptions are the studies by Bernhofen, El-Sahli, and Kneller (2016), who find large increases in bilateral trade between countries that have each installed one or more container-ready ports, and Gross (2017), who shows that converting 13,000 miles of U.S. railroad track to a standard gauge over a single

weekend in 1886 led to a sizable redistribution of traffic away from steamships. This paper describes how the UPC standard was created, provides empirical evidence on its diffusion, and shows how it influenced the nuts and bolts of trade in a wide variety of consumer goods.

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, revenue, 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 in the 1970s 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). Manufacturers were still motivated to participate, at least partly from fear that without an industry standard, each large grocery chain would require its suppliers to adopt a set of proprietary symbols (Brown, 1997, p. xv).⁴

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

³Although 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.

⁴Food 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 (Brown, 1997, p. 62).

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 revenue (Brown, 1997, pp. 119, 151).⁵ The second part of the code was assigned by the firm and could vary by product type, size, color, flavor, and other product characteristics. Computer code associated each prefix with a manufacturer, and each suffix with a product and a price.

Registering for a UPC is necessary but not sufficient to placing barcodes on products. It is the latter innovation that enables scanning by retail outlets. Printing the barcode symbol required manufacturers to redesign their product labels to make room for the symbol, and in some cases to invest in printing technologies that 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.

Adding UPC symbols to packaging only benefits downstream firms to the extent that they utilize scanners. At first, checkout-scanner adoption proceeded more slowly than the UPC registration process. For example, 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.” One year later, an editorial in *UPC Newsletter* noted that there had been just 78 retail scanner installations compared to 4,412 manufacturer UPC registrations (Uniform Product Code Council, 1976). One reason for this imbalance was that a single UPC registration — sufficient for a firm with 100,000 individual SKUs — was much cheaper than installing scanners at the checkout.⁶

Basker (2012, Figure 1) shows that scanner adoption began to accelerate around 1981.

⁵UPC adopters were not limited to a single registration, and many large firms registered for multiple prefixes. Thus, the registration cost does not appear to have been prohibitively high. In 1990, the number of digits in the prefix increased from five to six (Brown, 1997, p. 191).

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

By 1984, roughly eight percent of U.S. grocery stores had installed a scanner at checkout.⁷ Around that time, the major general-merchandise retailers started using scanners in the back end of stores for inventory management, often in conjunction with early EDI implementations. For example, Kmart reported that in 1981 it implemented back-end scanners whereby “store employees use a wand to scan hardline merchandise on the sales floor and in the stockroom, assuring accurate replacement of goods” (Kmart, 1982, p. 9). From 1982 to 1986, each of Walmart’s annual reports makes some reference to investments in UPC-based point-of-sale scanning systems. Abernathy and Volpe (2011) report that Kmart and Walmart required apparel suppliers to place a barcode on every item starting in 1983 and 1987, respectively.

3 Data

To study the diffusion and impact of the UPC, we construct a panel dataset containing information on UPC registrations, employment, and trademarking for approximately 779,300 manufacturing firms over the period 1975 to 1992, comprising 5.1 million firm-year observations.⁸ Table 1 reports summary statistics for this panel. On average, the firms in our data employed 72.8 persons and had \$14.4 million in annual revenue, though as with most firm-level datasets, the size distribution is highly skewed.⁹ There is considerable churn, with around 9.7 percent of active firms exiting the panel in any given year.

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

⁷The eight percent estimate is based on comparing 10,000 scanner installations to the 121,049 grocery establishments reported in the 1982 Census of Retail Trade. 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., and Beck, Grajek, and Wey (2011) study the adoption of scanners in Europe.

⁸In Appendix C, we present parallel results for a sample of 866,500 wholesalers over the same time period.

⁹We do not report medians for our sample, but reports from the Business Dynamics Statistics data indicate that the median-sized manufacturer in 1977 had between 10 and 19 employees. See http://www2.census.gov/ces/bds/firm/bds_f_szsic_release.xlsx.

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 flow of new U.S. registrations per year. After an initial spurt of registrations in 1974 and 1975, UPC adoption slowed for several years, before starting a steady climb that lasts through the end of our sample. We also observe some bunching in 1983, consistent with widespread adoption of the UPC by general-merchandise suppliers around that time. The right panel in Figure 1 shows the distribution of the size-class variable based on reported annual revenue (in millions of dollars) of the registering firm (Zimmerman, 1999, Appendix E). The vast majority of registrants are small firms with annual revenue under \$2 million.

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 Business Register (1975 to 1992). Details of the matching procedure are described in Appendix A.1. Ultimately, we successfully match between 40 and 50 percent of UPC registrations to establishments in three sectors: manufacturing, wholesale, and retail. The match rate is around 70 percent for firms with \$2 million or more in annual revenue, and around 40 percent for firms below that threshold.¹⁰

To create a panel data set, establishments are linked over time using the Longitudinal Business Database (LBD).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 Firm age is defined as the

¹⁰Figure A-1 in Appendix A.1 shows the match rate, by year and by firm size. A disproportionate share of early adopters were relatively large firms. For instance, the share of registrations in the \$0-2M size class increases sharply from 26 percent in 1972 to 86 percent in 1978, after which it stabilizes between 85 and 90 percent. For firms above \$2 million in revenue, our match rates are similar to those obtained by Jarmin (1999) for manufacturing plants and Kerr and Fu (2008) for patent filers.

¹¹The LBD is described in detail in Jarmin and Miranda (2002).

¹²Details of the aggregation procedure are described in Appendix A.2.

difference between the current year and the year that the firm identifier first appears.¹³ Table 1 shows that 3.8 percent of the observations in our sample belong to a manufacturer that registered for a UPC by 1992. At the firm-year level, the UPC adoption rate is 1.9 percent, suggesting that adopters held a UPC for about half of the years when they appear in the sample. These numbers understate the rate diffusion of the UPC because of the highly skewed firm-size distribution, and the fact that larger firms were more likely to adopt early, as we show in the next section.

Our analysis of network effects relies on two measures of aggregate UPC adoption to proxy for the installed base of scanners. The first variable is based on UPC adoption by rivals in the same four-digit SIC code as a focal firm, and is denoted by $\overline{\text{UPC}}_{it}$. The second variable measures UPC adoption by manufacturers in *other* four-digit SIC codes that sell through the same retail channels as a focal firm, and is represented by $\widehat{\text{UPC}}_{it}$. For example, consider a firm in SIC 2033, “Canned fruits and vegetables.”¹⁴ For this firm, $\overline{\text{UPC}}_{it}$ measures the employment-weighted UPC adoption rate of other firms in SIC 2033, whereas $\widehat{\text{UPC}}_{it}$ reflects the employment-weighted UPC adoption rate of firms in other industries, such as salad-dressing producers, tobacco producers, and magazine publishers, that also sell to grocery stores. We briefly describe how the two variables are created here, and provide more detail in Appendix A.3.

Because firms may have one or more establishments (plants, warehouses, etc.), each with a different SIC, we calculate $\overline{\text{UPC}}_{it}$ 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 excluding the focal firm), and the second

¹³The first year is either 1972, if any of the firm’s establishments appear in the 1972 CM, or 1975 and later, because the LBD starts in 1975.

¹⁴As detailed in Appendix A.2, firms are classified by their predominant industry. Firms that operate multiple establishments are therefore classified in a single industry despite having some establishments in other industries.

step aggregates this measure up to the firm level. Specifically, we calculate

$$\overline{\text{UPC}}_{it} = \sum_{e \in E(i)} w_e \left[\sum_{k \in M(e)} z_k \text{UPC}_{kt} \right] \quad (1)$$

where $E(i)$ denotes all establishments at firm i ; w_e is establishment e 's share of total employment at firm i in year t ; the set $M(e)$ contains all establishments in the same four-digit SIC code as establishment e , excluding any establishments owned by firm i ; z_k is establishment k 's share of total employment within the set $M(e)$ in year t ; and UPC_{kt} is an indicator for UPC adoption at establishment k by year t .

To create $\widehat{\text{UPC}}_{it}$ we rely on data from the Census of Retail Trade (CRT), along with a hand-made concordance between four-digit manufacturing SIC codes and ‘‘Broad Line’’ product categories in the CRT data (examples of broad lines are food, women’s apparel, and furniture). This concordance allows us to compute a pair of weights: u_{rm} measures the share of retail industry r 's 1977 revenue derived from manufacturing industry m 's products, and s_{mr} measures the share of manufacturing industry m 's 1977 sales through each retail channel r . For each firm i in manufacturing industry j , we then compute

$$\widehat{\text{UPC}}_{it} = \sum_{r \in R} s_{jr} \left[\sum_{m \in \{M \setminus j\}} u_{rm} \overline{\text{UPC}}_{mt} \right] \quad (2)$$

where R is the set of all 4-digit retail SIC codes; $\{M \setminus j\}$ is the set of all manufacturing industries except for j ; and $\overline{\text{UPC}}_{mt}$ is the employment-weighted industry average UPC adoption for manufacturing industry m in year t .¹⁵ Table 1 shows that $\overline{\text{UPC}}_{it}$ averages 16.1 percent and $\widehat{\text{UPC}}_{it}$ averages 7.1 percent across all manufacturing firm-years in our panel. These numbers are substantially higher than the firm-level UPC adoption rate because of the employment weights and the fact that large firms adopted the UPC earlier.

¹⁵In this calculation, we treat each firm as having a predominant manufacturing SIC code, instead of aggregating up using establishment weights as we did for UPC_{it} .

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 (forthcoming).¹⁶ 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 an indicator that a firm applied for at least one new TM that eventually became a registered TM.¹⁷ To avoid double counting TMs that change hands, we restrict our counts to the original applicant. Table 1 shows that 8.1 percent of the observations in our panel belong to a firm that applied for at least one new trademark and the annual probability of filing a new TM was 1.6 percent, which together imply that trademarking firms registered a new TM around once every five years.

Finally, we create 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 industry-level UPC adoption, calculated from the micro data, 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 or only imports from Canada, with closely related industries (to avoid zero cells when we take logs), this yields a strongly

¹⁶Matching 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.

¹⁷This 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.

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 in Figure 2. 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.

Our data 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. General-merchandise stores soon “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 3 reinforces the idea that the UPC was widely adopted within the grocery supply chain before spreading to general merchandise. Each panel plots UPC adopters’ share of firms and employment within six selected manufacturing industries. All panels are on the same scale, but the firm share and employment share use different axes.

¹⁸Examples 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).

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

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 of UPC registrants, however, remains fairly stable at 60 percent, reflecting the fact that large firms registered early and later registrants are small.²⁰ In chemical production, which includes pharmaceuticals, adoption by large firms occurs early but both the employment share and the firm share of adopters increase steadily over time. Both food and chemical manufacturers are likely to sell through the grocery supply chain. The other four industries in Figure 3 (apparel, electronics, furniture, and textile manufacturing) mostly supply their respective specialty 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 growth because here, too, larger firms adopt earlier.

4.2 Network Effects

The UPC is a classic case study for two-sided network effects. The 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, computer hardware and software, and 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).

²⁰Food 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.

With comprehensive data on UPC adoption, scanner adoption, and supply-chain links, one could estimate network effects via a system of equations

$$\Delta\text{UPC}_{it} = \alpha^u + \beta^u \text{Scanner}_{J(i),t} + \varepsilon_{it}^u \quad (3)$$

$$\Delta\text{Scanner}_{jt} = \alpha^s + \beta^s \text{UPC}_{I(j),t} + \varepsilon_{jt}^s \quad (4)$$

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 $\text{Scanner}_{J(i),t}$ and $\text{UPC}_{I(j),t}$ measure the stock of scanning retailers $J(i)$ or bar-coding manufacturers $I(j)$ within the focal firm's supply chain.²¹ Indirect network effects imply positive values for both β^s and β^u .

Estimating this model directly is not possible with our data. In particular, we have only coarse industry-level information about supply chains, and in most cases no data on scanner adoption.²² To derive a feasible reduced-form specification, we integrate Equation (4) over retailers k and time periods τ to obtain

$$\text{Scanner}_{J(i),t} = \sum_{\tau \leq t} \sum_{k \in J(i,\tau)} \alpha^s + \beta^s \text{UPC}_{I(k),\tau} + \varepsilon_{k\tau}^s \quad (5)$$

where $J(i,\tau)$ is the set of retailers supplied by τ by manufacturer i that have not adopted scanning by τ . The main message of Equation (5) is that the stock of scanning retailers in manufacturer i 's supply chain can be expressed as a weighted average of UPC adoption by other manufacturers that supply those same retailers. Substituting Equation (5) into Equ-

²¹In a fully structural dynamic model, the key explanatory variables would likely be written as $\mathbb{E}_t(\text{Scanner}_{J(i),t+1})$ and $\mathbb{E}_t(\text{UPC}_{I(j),t+1})$, to denote the *expected* future stock of barcodes or scanners. We adopt a linear specification and ignore questions of how to model expectations of future adoption for simplicity of exposition.

²²Appendix B.1 provides an analysis of grocery-store scanner adoption between 1974 and 1984, using data from Basker (2012). Although this exercise lends some plausibility to the indirect-network-effect interpretation of our main results, we cannot estimate a two-sided model for the larger data set because we lack scanner adoption outside the grocery sector and for later years.

tion (3) suggests that we can estimate $\beta^s \times \beta^u$ using variation in UPC adoption by other manufacturers that sell through the same channels as manufacturer i . This reduced-form parameter should be positive when there are indirect network effects.

In practice, we replace $\text{Scanner}_{J(i),t}$ in Equation (3) with either $\widehat{\text{UPC}}_{it}$ or $\overline{\text{UPC}}_{it}$ to obtain the reduced-form specification

$$\Delta \text{UPC}_{it} = \lambda_{at} + \beta \widehat{\text{UPC}}_{it} + X_{it}\theta + \varepsilon_{it} \quad (6)$$

where λ_{at} are a full set of firm-age by calendar-year fixed effects, and X_{it} are exogenous controls.²³ 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 hazard of UPC adoption if all other manufacturers selling through the same retail channels switched from being non-adopters to adopters.²⁴

Manski (1993) discusses identification of models such as Equation (6), where an individual choice is regressed on a group average of the same outcome. We interpret β as what Manski calls a correlated effect: manufacturers with similar values of $\widehat{\text{UPC}}_{it}$ (or $\overline{\text{UPC}}_{it}$) make correlated choices because they face a similar institutional environment, specifically, downstream customers that have installed scanners. The alternative interpretation, which Manski calls an endogenous effect, is that UPC adoption by other manufacturers has a direct causal impact on the decisions of a focal firm. Empirical models of indirect network effects typically rule out endogenous effects, which are also called direct network effects or “same side” externalities, to achieve identification (Rysman, 2019). This is a plausible assumption in our setting, where spillovers among upstream UPC adopters are likely to be minimal in the absence of downstream scanning.

Table 2 reports estimates based on Equation (6). Standard errors are clustered by

²³Appendix D shows how to derive this reduced-form specification, up to a monotonic transformation of $\widehat{\text{UPC}}_{it}$, from a linear system of first-order differential equations analogous to Equations (3) and (4).

²⁴Jenkins (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.

four-digit firm SIC, and all models include controls for lagged firm employment and vertical integration (i.e., an indicator for firms with one or more wholesale or retail establishments). The first two columns present estimates from a pure correlated-effects model, which implicitly assumes no unobserved industry-level heterogeneity. The network-effect parameter is positive and statistically significant for both channel ($\widehat{\text{UPC}}_{it}$) and rival ($\overline{\text{UPC}}_{it}$) adoption. To provide a sense of the economic magnitudes, note that a one-standard-deviation change in $\widehat{\text{UPC}}_{it}$ (as reported in Table 1) doubles the baseline hazard of UPC adoption (from 0.34 to 0.67 percent per year), and a one-standard-deviation change in $\overline{\text{UPC}}_{it}$ increases the baseline hazard by approximately 130 percent.

One concern with the initial estimates in Table 2 is that correlated industry-level unobservables, such as a lower average cost of UPC adoption, might be mistaken for network effects. We find the indirect network-effects interpretation more plausible because the costs of UPC adoption do not seem to have a large industry-specific component. It is also reassuring that the coefficients on $\widehat{\text{UPC}}_{it}$ and $\overline{\text{UPC}}_{it}$ are quite similar, since the latter measure excludes any variation produced by UPC adoption in the same four-digit manufacturing industry as a focal firm. Nevertheless, to address the possibility of industry-level omitted variables, such as a coordinated effort to start UPC and scanner adoption in a particular industry, the third and fourth columns in Table 2 present results from a specification that controls for industry fixed effects and time-varying log industry value-added. For both channel and rival adoption, the indirect network effect parameter remains positive and statistically significant. Although point estimates decline by 30 to 50 percent relative to the first two columns, in the case of $\widehat{\text{UPC}}_{it}$ the confidence intervals include a wide overlap range.²⁵

²⁵One concern with the inclusion of SIC fixed effects is that both $\widehat{\text{UPC}}_{it}$ and $\overline{\text{UPC}}_{it}$, which aggregate prior UPC adoption decisions, implicitly contain lagged outcomes, leading to a violation of strict exogeneity. The resulting bias is likely to be small, however, because the time (firm-year) dimension of our panel is large. In particular, Nickell (1981) shows that under the assumption of sequential exogeneity, the bias of the within estimator is inversely proportional to T . We cluster at the four-digit SIC level, which implies a panel of around 600 manufacturing industries each containing an average of $T = 8,500$ observations.

Finally, to check whether our results actually reflect manufacturers imitating one another, or perhaps a growing awareness of UPC, we conduct a geographic placebo test. Specifically, we re-define $M(e)$ as the set of establishments that share a three-digit ZIP code with establishment e but are owned by a different firm, re-calculate $\overline{\text{UPC}}_{it}$ according to Equation (1), and re-estimate Equation (6). If spillovers in UPC adoption are driven by imitation or awareness, we would expect a significant geographic component.²⁶ The estimate in the last column of Table 2 shows that geographic spillovers in UPC adoption are negligible. It remains possible that firms imitate geographically distant direct competitors, but not other nearby firms. In Section 5.1, however, we show that firm-level employment increases following UPC adoption, providing further evidence that diffusion is driven by vertical interactions rather than pure imitation.

5 Impacts of UPC Adoption

This section estimates the impact of UPC adoption on several outcomes. At the firm level, we analyze employment, revenue, 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 and Revenue

Difference-in-Difference Estimates

To estimate the impacts of UPC adoption, we use the following difference-in-difference specification:

$$Y_{it} = \alpha_i + \gamma_{mt} + \lambda_{at} + \beta \text{UPC}_{it} + \varepsilon_{it} \quad (7)$$

²⁶Hillberry and Hummels (2008) use the Commodity Flow Survey to show that shipments are highly localized, so geographic correlations in UPC adoption might still capture the impact of supply-chain relationships. The concentration of shipments is less pronounced at the three-digit ZIP-code level than at the narrower five-digit level, however, and also implies that our placebo test is conservative.

where Y_{it} is firm i 's logged employment or revenue, or an indicator for trademark registration status in year t ; α_i is a firm fixed effect; γ_{mt} is an industry-year effect linked to the predominant industry of firm i ; λ_{at} is a firm-age by calendar year effect; and UPC_{it} is an indicator that turns on if and only if firm i registered for a UPC by year t . The industry-year fixed effects provide a flexible specification of the outcome's dynamics in each of the four-digit SIC manufacturing industries in our sample. The age-year 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 construct two different estimation samples for this analysis. 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.²⁷ The one-to-one matched sample is 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.²⁸ For firms aged two through five at registration, we match using vintage, year $(t - 1)$ employment, and log employment growth between year of birth and year $(t - 1)$.²⁹ Finally, we match firms aged six and over at the time of registration to

²⁷Employment and revenue are both proxies for firm size. We do not match on revenue or revenue growth because revenue is available only in five-year intervals from the Economic Census. Instead, we report estimates from the employment-matched sample using total revenue as the outcome variable.

²⁸We 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.

²⁹We find the closest match on employment growth, with the restriction that the two firms' employment growth cannot differ by more than 0.5 percent.

other firms that are at least six 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. Matching is done without replacement.

We do not restrict matches to have the same manufacturing industry SIC for several reasons. First, if UPC adoption is driven by the downstream demand for barcodes, which varies more between industries than within them, matching across industries should reduce concerns about selection on firm-level unobservables. Intuitively, the experiment we would like to run randomly assigns manufacturing firms to supply chains with and without downstream scanners. Between-industry matching brings us closer to this notional experiment, whereas within-industry matching raises questions about self-selection, given that adopters and non-adopters in the same industry are exposed to similar supply chains. Second, contamination is a concern with intra-industry matching: controls may be affected by their competitors' adoption 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.

In the matched-sample analysis, counterfactual outcomes for UPC adopters are estimated by actual outcomes at similarly sized non-adopters that exhibit similar pre-adoption trends (relative to an industry-specific baseline) over the same time period. To account for staggered adoption, the matched sample regressions also include a post-adoption indicator that turns on for each non-adopter (i.e., control) firm in all years following the UPC registration by its matched adopter (treatment) firm. This post-adoption indicator variable ensures that β is identified by differences between adopters and non-adopters in the post-adoption time-period. We interpret the matched-sample estimates of β as an average treatment effect for treated firms.

Table 3 reports coefficient estimates based on Equation (7). The coefficient on UPC adoption is positive and statistically significant in all specifications, and the magnitude of the estimates is quite similar for the employment and revenue outcomes. The baseline OLS

estimates imply a 16 percent increase in employment and a 20 percent increase in revenue following UPC adoption, whereas the matched sample estimates suggest a 13 percent increase in employment and a ten percent increase in revenue.³⁰

The similar difference-in-difference estimates for revenue and employment suggest that UPC adoption influenced firm size more than productivity or the labor share. Because Total Factor Productivity (TFP) is typically measured at the establishment level and is not available for all establishments in all years, it is difficult to obtain directly comparable estimates. Nevertheless, Appendix B.3 reports estimates from an establishment-level version of Equation (7) for establishments in the Census of Manufactures and the Annual Survey of Manufactures, using revenue-based TFP from Foster, Grim, and Haltiwanger (2016) as the outcome. We find no evidence of a relationship between UPC adoption and manufacturing TFP — the coefficient on UPC adoption is uniformly small and statistically insignificant. This null result is consistent with the idea that even though upstream adoption of UPC is a necessary condition for bar-coding to work, the majority of productivity gains accrued to downstream retailers that invested in scanners and other complementary technology. And indeed, Basker (2012) estimates that labor productivity in grocery stores increased by 4.5 percent in the initial years following a scanner installation.

Event Study

Even with matching, it is hard to say to what extent the regressions in Table 3 estimate a selection effect as opposed to a causal effect of UPC adoption. To get a better handle on this question, we estimate a series of event-study regressions using employment as the outcome

³⁰One potential concern with this specification is that survival rates could differ for adopters and non-adopters. We have estimated the matched-sample difference-in-difference regressions on a balanced sample that drops *both* the adopter and its matched control when either firm exits the sample, and confirmed that this produces qualitatively similar estimates. In Appendix B.2, we report hazard models showing that UPC adoption is positively correlated with survival.

variable.³¹ Our main specification is:

$$\ln(\text{Employment}_{it}) = \alpha_i + \gamma_{mt} + \lambda_{at} + \sum_{\tau=-6}^{12} (\delta_{\tau} + \beta_{\tau}\text{UPC}_i) + \varepsilon_{it} \quad (8)$$

where α_i , γ_{mt} , and λ_{at} are defined above; δ_{τ} is a vector of indicators that turn on only if either an adopter or its matched control registered for a UPC in year $(t+\tau)$; and β_{τ} measures a treatment effect at τ years before or after adoption. We use a single indicator for $\tau \leq -6$ and normalize $\delta_{-1} = \beta_{-1} = 0$. To ensure that we do not include future adopters in the control group, we restrict this regression to observations in 1986 and prior years.³²

Figure 4 plots the event-study coefficients, β_{τ} , for the matched sample.³³ The connected dots correspond to point estimates, and the error bars are upper and lower 95 percent confidence limits. By construction, relative employment of adopters and non-adopters between $(t-5)$ and $(t-1)$ is nearly identical and statistically indistinguishable. Following adoption, the treated and control firms clearly diverge: employment increases by about ten percent in the year of adoption, and then by another ten percent over the next two years. The abrupt increase in relative employment in the year of UPC adoption is a striking result. The discrete jump suggests to us that manufacturers adopted the UPC specifically to integrate with retail supply chains. This does not mean 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, and not merely a proxy for adopting UPC-related technologies and business practices, which would produce more gradual and consistent short-term employment growth. Although the confidence intervals

³¹It is not possible to provide event-study estimates for the revenue outcome, because revenue is only available in five-year intervals.

³²The coefficients δ_{τ} captures common trends in the treatment and control firms' employment before and after adoption of the UPC by the treatment firms.

³³The figure omits β_{11} and β_{12} , which tend to be imprecisely estimated due to small cells, raising disclosure concerns.

increase over time, the point estimates imply that growth continues at least 5–7 years after UPC adoption. This gradual increase in employment (relative to the counterfactual) in later years is consistent with the idea that UPC registration is correlated with downstream scanner adoption, along with a broader set of technological and organizational changes linked to supply-chain automation.

Figure 5 provides two points of comparison that assist in the interpretation of the matched sample event-study. First, Figure 5(a) shows event-study coefficients for the full sample.³⁴ Consistent with the results in Table 3, we observe a strong selection effect: in the years prior to registration, soon-to-adopt firms grow faster than controls. This raises a concern that UPC adoption is correlated with some combination of unobserved managerial ability and opportunity. In principle, we have addressed this concern by matching on firm growth, and by including both firm and industry-year fixed effects in Equation (8). The absence of any measurable impact of UPC adoption on TFP (shown in Appendix B.3) also suggests that there is little selection along this margin. Nevertheless, it would be reassuring to see that fast-growing non-adopters in the same industries as UPC adopters do not experience any “UPC adoption” effect. Figure 5(b) presents results from that type of placebo test.

To construct the sample used in the Figure 5(b), we first match each UPC adopter to a single non-adopter in the same two-digit SIC, using the employment-level and -growth matching procedure described above. We then discard the UPC adopters, and treat the remaining sample of matched controls *as if* they had adopted UPC in the same year as their discarded twin. Finally, we match each firm in this placebo-adopter sample to its own control (in this case allowing for between-industry matching, as we do for the matched sample), and re-estimate the event-study model of Equation (8). The coefficients graphed in Figure 5(b) show that non-adopters from the same broad industries as the UPC adopters, having similar

³⁴In the regression that produced this figure, we omit the δ_τ coefficients because “years to adoption” τ is not defined for control observations in the absence of a matching procedure.

pre-adoption size and growth trends, do not exhibit any meaningful treatment effect. This lends additional confidence to our preferred interpretation of the matched-sample results: a casual impact of joining scanner-enabled supply chains.

Network Effects Revisited

To provide more evidence on the role of network effects, we can extend the baseline difference-in-difference model to ask whether the impact of UPC adoption on employment and revenue increases with downstream scanner adoption. As in the diffusion analysis, $\widehat{\text{UPC}}_{it}$ and $\overline{\text{UPC}}_{it}$ serve as our proxies for downstream scanner adoption. To implement a test for network effects, we interact one of these proxies with an indicator for focal-firm adoption and add it to the difference-in-difference specification in Equation (7). Specifically, we estimate

$$\ln(Y_{it}) = \alpha_i + \gamma_{mt} + \lambda_{at} + \beta \text{UPC}_{it} + \delta \text{UPC}_{it} \widehat{\text{UPC}}_{it} + \varepsilon_{it} \quad (9)$$

where Y is either employment or revenue. In this specification, the main effect of $\widehat{\text{UPC}}_{it}$ is absorbed by the industry-year fixed effects, and we include an indicator for all post-adoption years (for both UPC adopters and matched controls) as described above.³⁵ All of our results are based on the matched sample.

Table 4 reports estimates of the direct effect of adoption, β , and the interaction term, δ , for both proxies for scanner adoption. The main effect of UPC adoption, or equivalently the impact of UPC adoption for the first adopter in an industry, is reported in the first row of the table. This coefficient estimate is positive and statistically significant in all models, and indicates either a 9–10 percent increase in employment or a 5–7 percent increase in revenue.

The interaction term, which we interpret as a measure of network effects, is also positive

³⁵For regressions that use $\overline{\text{UPC}}_{it}$ to proxy for scanning, the main effect is not precisely co-linear with the industry-year dummies because the focal firm is omitted from the “industry average.” In practice, we include a main effect for $\overline{\text{UPC}}_{it}$ in the regressions reported below.

and statistically significant across all models.³⁶ One way to interpret the economic significance of δ is to note that a one-standard-deviation increase in $\widehat{\text{UPC}}_{it}$ raises the predicted marginal effect of UPC adoption from ten to 12 percent for employment, and from 7 to ten percent for revenue.

In this model, we interpret both rival and channel UPC adoption as proxies for downstream scanner adoption.³⁷ Under that interpretation, our results show that scanner adoption by downstream customers amplifies the impact of UPC adoption on manufacturing firm size. Basker (2012) provides a complementary result for the retail side of the UPC platform: during a period when barcoding variable-weight products, such as fresh produce, was relatively rare, grocery stores that sold more packaged goods realized greater labor productivity gains from scanner adoption.

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, “The diffusion throughout the Food and Beverage sector has been steady with associated product proliferation, much larger stores and the addition of numerous new departments and an approach to the early objective of one-stop shopping.” We investigate this hypothesis using registered TM applications as a proxy for variety-increasing product innovation.

³⁶We have also estimated these regressions using a one-year lag of competitors’ adoption, $\overline{\text{UPC}}_{i,t-1}$. Results are very similar. In a previous draft, as a robustness check, we estimated a similar model (with firm-age and year, but not industry-year, fixed effects) on a matched sample within a four-digit firm SIC. This restriction reduces the sample size by 40 percent, and although the coefficient estimates change very little, estimates of the interaction effect, δ , lose statistical significance due to larger standard errors.

³⁷To check whether the interactions terms in Table 4 might be picking up treatment heterogeneity by firm size, given that larger firms tended to adopt earlier when $\widehat{\text{UPC}}_{it}$ was smaller, we estimated a model that allowed the effect of UPC adoption on employment to vary by firm-size quartile. In this model, we found no clear relationship between firm size and the size of the coefficient on UPC_{it} .

As a starting point, Figure 6 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*.³⁸ Both series are “ocular reproductions” of data reported in Sullivan (1997).³⁹ The dotted line is a count of new TM applications for grocery-related products that we constructed from the USPTO data.⁴⁰

Figure 6 helps motivate a firm-level trademark analysis in two ways. First, it shows that TM applications are strongly correlated with 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 3. 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 difference-in-difference specification of Equation (7). The outcome variable is an indicator that turns on if firm i files for one or more new trademarks in year t . To address the selection effects observed above, we use prior TM registrations to create a matched sample. Each UPC adopter in the matched sample has a unique control. 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 registra-

³⁸New 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).”

³⁹We 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.

⁴⁰In order to restrict our count of TM applications to the grocery industry, we focus on applications with one or more three-digit Nice codes corresponding to food, beverages, pharmaceuticals, or paper products. We adjust for missing data in the years before 1977 using a procedure described in Appendix A.4.

tion, 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 , matched by the cumulative number of TMs they have registered between years $(t - 5)$ and $(t - 1)$. Results are presented in Table 5.⁴¹

Estimates for both the full and the matched sample show a statistically significant 4.5 percentage point increase in the probability of trademarking following UPC registration. This effect is large relative to the 1.6 percent baseline probability of filing a TM, but appears reasonable compared to the 20 percent annual filing probability for firms that registered at least one new TM during the sample period.

We also estimate an event-study specification for trademarking, based on Equation (8) and using the matched sample to address potential selection effects. Figure 7 graphs the β_τ coefficients. By design, there is no pre-registration trend difference between adopters and matched controls. The probability of TM registration increases steadily in the decade after UPC adoption; ten years out, the probability of a TM registration is 13 percentage points higher than the counterfactual rate.

To summarize our firm-level analyses, we find that manufacturer UPC registration is associated with economically and statistically significant increases in employment, revenue, and trademark registrations, whereas we find no relationship between UPC adoption and revenue-based TFP. Interpreting these results requires care. Although we use matching to remove selection effects, and estimate a placebo event study to show that even fast-growing non-adopters from the same manufacturing industries do not exhibit similar outcomes, 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 were in place, upstream UPC adoption helped manufac-

⁴¹The number of observations in the matched regressions differs from the number of observations in the matched regressions from Tables 3 and 4 because the matching procedure is different.

turers achieve scale by supplying large retailers. The trademark results suggest that joint adoption of scanning and barcodes 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 the scale of modern retail chains. If scanners and barcodes lower retailers' cost of managing a large assortment of goods, then imports are one channel through which they could obtain greater variety, complementing other channels such as adding domestic suppliers and exerting demand-side pressure on existing suppliers to increase their product offerings.⁴² The UPC-registration data allow us examine this hypothesis by measuring the industry-level association between domestic retail technology adoption and international trade.

The outcome variables in our trade analysis are log U.S. imports and import penetration, measured at the manufacturing industry-year level.⁴³ Our estimates are based on the

⁴²Basker and Van (2006) offer a formal model of another potential channel linking UPC adoption to trade: 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.

⁴³Import penetration is defined as the ratio: Imports / (Shipments + Imports - Exports).

following reduced-form specification:

$$\text{Trade}_{mt} = \alpha_m + \lambda_t + \beta \widehat{\text{UPC}}_{mt} + X_{mt}\theta + \varepsilon_{mt} \quad (10)$$

where Trade_{mt} measures log imports or import penetration for industry m in year t ; $\widehat{\text{UPC}}_{mt}$ is employment-weighted domestic industry UPC adoption; α_m are industry fixed effects; λ_t are calendar-year fixed effects; and the error term ε_{mt} is clustered at the industry level. We also include time-varying industry-level controls, X_{mt} , for log industry value-added, log capital-labor ratio, and the log ratio of production to non-production workers.

Our main explanatory variable, $\widehat{\text{UPC}}_{mt}$ (alternatively $\overline{\text{UPC}}_{mt}$), is based on adoption by domestic manufacturers. Although we do observe some UPC registrations with a foreign address, it is unlikely that they reflect the full extent of foreign adoption, given that domestic firms can register for a UPC and have international suppliers print that domestic code on their packaging.⁴⁴ More importantly, the domestic registration data are well-suited to our purposes, because they can be mapped onto a particular manufacturing industry. Given our previous results providing evidence of network effects in the diffusion process, we 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 increase retailers' benefits from, or reduce their 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 expand their selection, (c) an output-expanding effect if retailers pass through lower prices to consumers, and (d) increased imports of intermediate goods as foreign manufactur-

⁴⁴For example, there was a large increase in Asian registrations in 1983 — the year when Kmart began to require that all apparel suppliers use barcodes.

ers begin to supply components to domestic producers. It is less clear what we should expect for exporting. Because the UPC is a domestic standard, it is tempting to view exports as a placebo test. In practice, our estimates of the relationship between domestic UPC adoption and exports are statistically insignificant and close to zero in almost all specifications, so we focus on the import regressions.

Table 6 presents the results of our trade regressions. Across all four models, we find a statistically significant positive relationship between domestic UPC adoption and imports. The magnitude of the coefficients implies that a one-standard-deviation increase in industry-level UPC adoption is associated with a 6–7 percent increase in imports.⁴⁵ This result is robust to dropping Canadian imports (which might be influenced by the 1988 Canada-U.S. Free Trade Agreement) and also two different approaches to excluding trade in intermediate goods.⁴⁶

There are several reasons to be cautious about the trade results. In one sense, it is clear that the estimates are not causal: adding numeric labels to domestic producers' labels should not cause an increase in imports. 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) could stimulate

⁴⁵The within-industry standard deviation of $\overline{\text{UPC}}_{mt}$ is approximately 0.125.

⁴⁶To exclude intermediates, we limited the estimation sample to the set of manufacturing industries in our CRT-SIC concordance described in Appendix A.3, which clearly sell some products through retail channels. 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). These results are not reported but are qualitatively similar to reported results.

domestic technology adoption.

In spite of these concerns, we believe these regressions provide some of the first evidence linking import growth directly to changes in retail technology. Moreover, the correlation 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; Dunlop, 2001) and subsequent increases in industry concentration (e.g., Basker, Klimek, and Van, 2012; Hsieh and Rossi-Hansberg, 2019). This paper is the first to measure the effects of UPC adoption on upstream employment, revenue, product innovation, and industry-level imports, providing a natural complement to the literature on retail productivity (e.g., Foster, Haltiwanger, and Krizan, 2006; Basker, 2012) and a new addition to the empirical literature on the effects of industry standards.

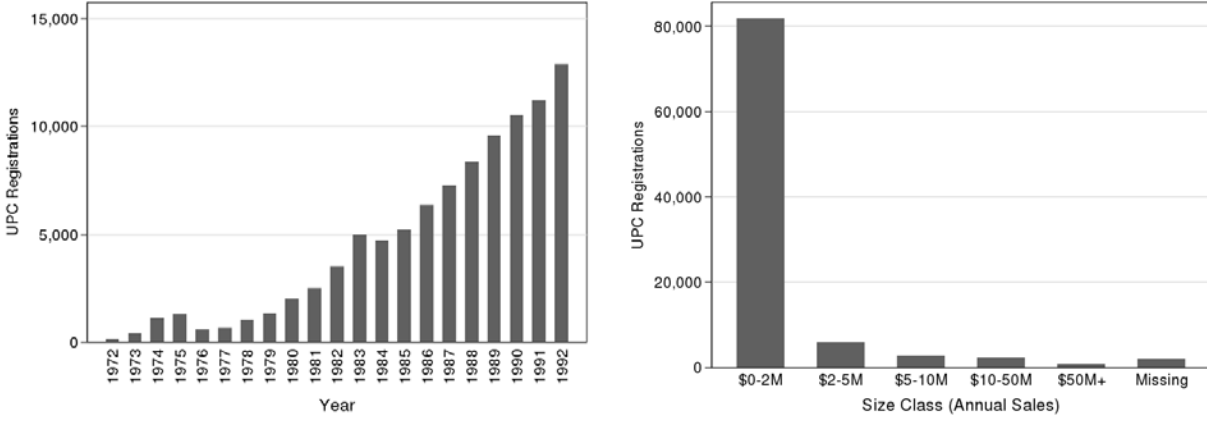
We show that early UPC adoption is strongly correlated with firm size and that the timing of UPC adoption varied across industries. Many large food-and-drug manufacturers had already adopted the UPC by the mid 1970s, whereas adoption by apparel, furniture, and textile manufacturers remained at low levels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

Our investigation of the impacts of UPC adoption suggests that both upstream and downstream firms benefited on several margins. For manufacturers, we find that both revenue and employment increased with and following adoption, consistent with receiving larger orders from retailers. The timing of the employment effects revealed by our matched-sample event-study regressions suggests that UPC adoption is associated with business diversion, whereby manufacturers integrate into the supply chain of large downstream retailers. These findings help explain why manufacturers embraced the UPC, even if the benefits were expected to accrue mostly to retailers. Early adopters sought to prevent standards fragmentation, whereby each large retailer would require its suppliers to use a proprietary symbol. Later on, particularly after Kmart and Walmart required their suppliers to barcode all items, the pressure from the demand side became explicit. Our results show that manufacturers benefited from UPC adoption, albeit indirectly, through increased orders; and that these benefits grew as UPC proliferated through the retail channel.

The downstream effects of UPC adoption are harder to assess. Retail productivity growth presumably reflects the direct benefits of scanning, as well as the increased scale and scope made possible by UPC and complementary technologies. Because we do not have explicit data on buyer-supplier relationships, 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 is disproportionately due to firms that registered for a 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.

Figure 1. New UPC Registrations



(a) By Year

(b) By Size Class

Figure 2. UPC Diffusion by Firm Revenue

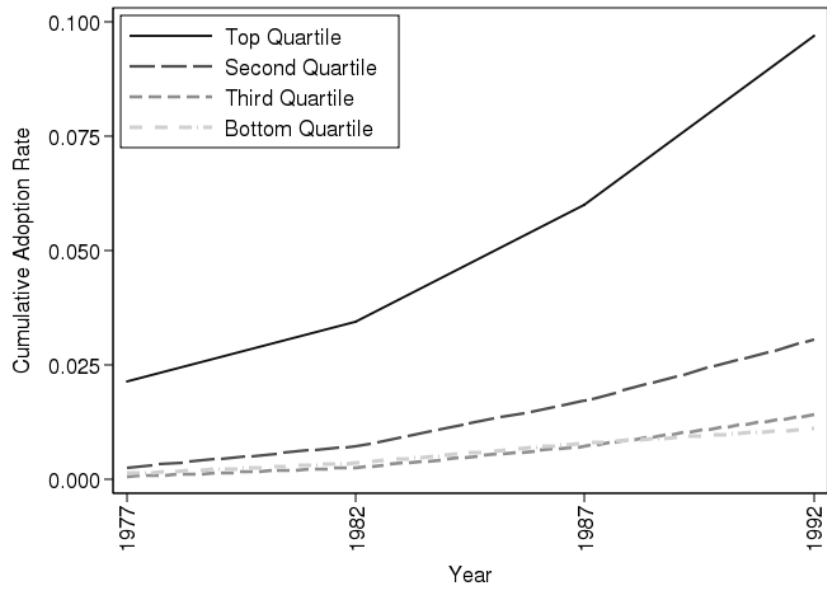


Figure 3. UPC Diffusion for Selected Two-Digit SIC Manufacturing Industries

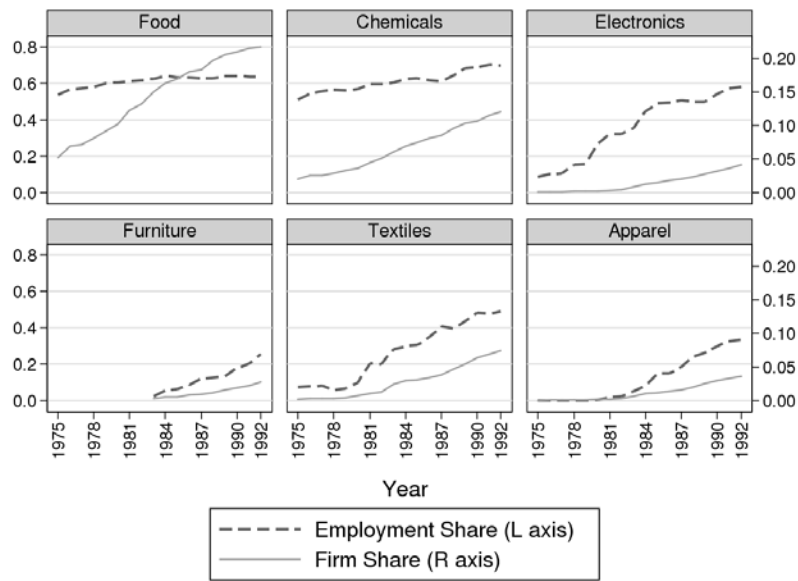
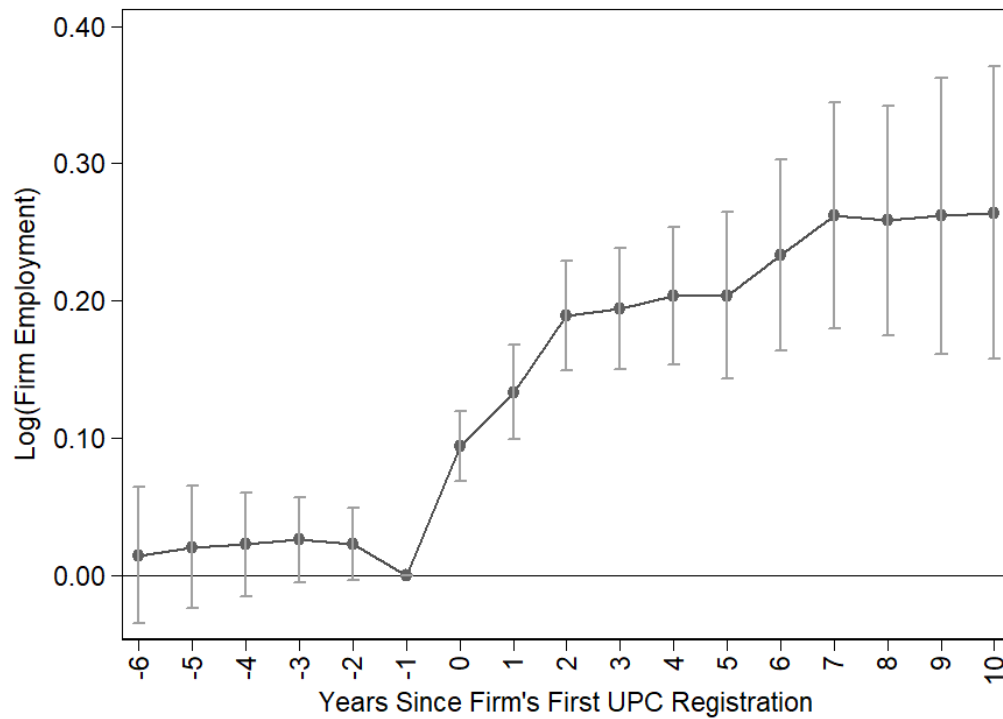
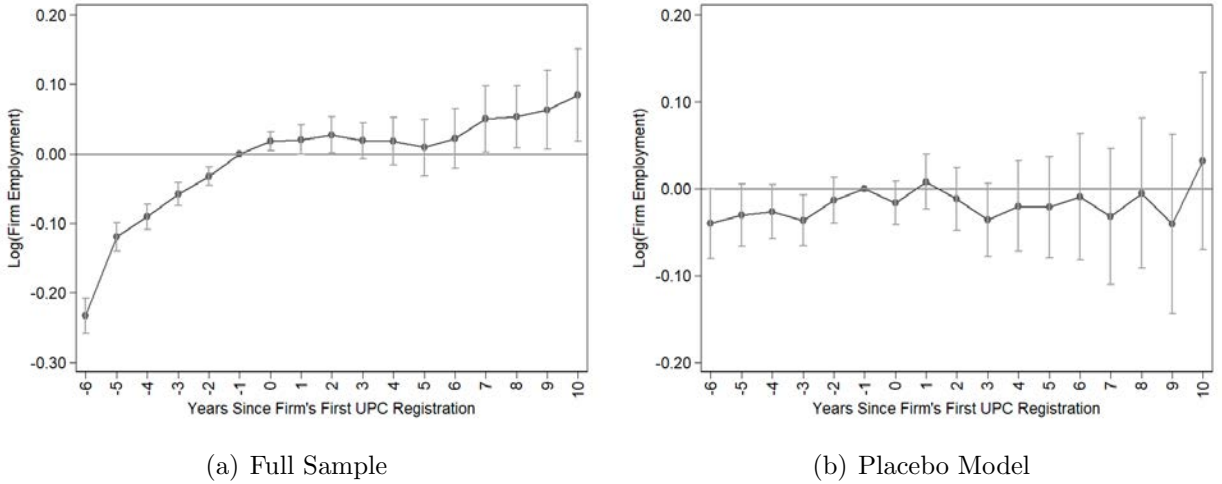


Figure 4. Matched Sample Event Study: Employment



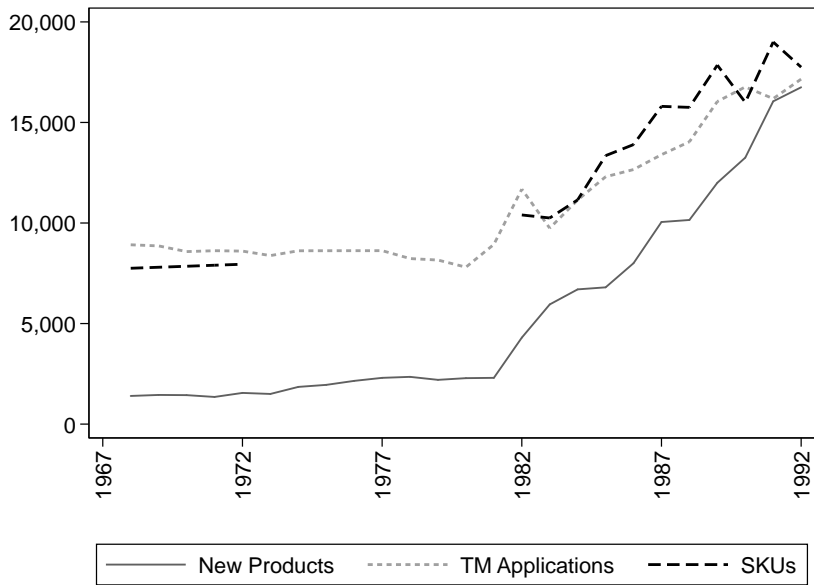
Note: Vertical bars represent 95 percent confidence intervals

Figure 5. Full Sample and Placebo Event Studies: Employment



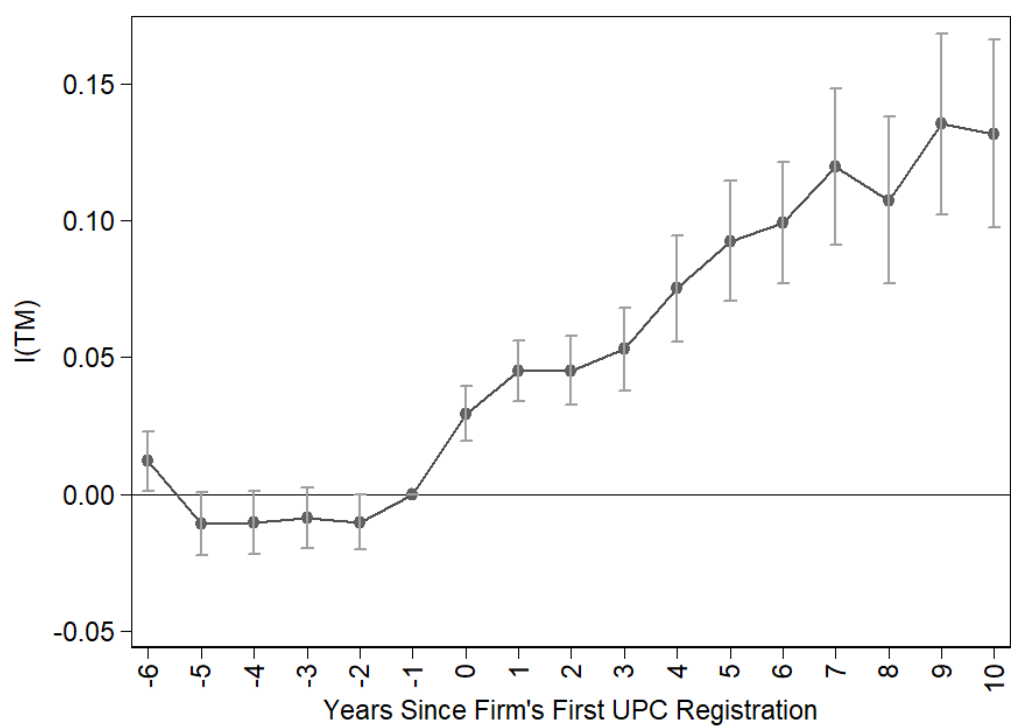
Note: Vertical bars represent 95 percent confidence intervals

Figure 6. 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 Appendix A.4 for details).

Figure 7. Matched Sample Event Study: Trademark Registrations



Note: Vertical bars represent 95 percent confidence intervals

Table 1. Summary Statistics

	Mean	SD
Employees	72.8	1,540
Revenue ^a	14.4	0.592
$\mathbb{I}[\text{Exit} \text{Alive}_{t-1}]$	0.097	0.296
UPC adoption: UPC_{it}	0.019	0.135
UPC adopter	0.038	0.190
Channel adoption: $\widehat{\text{UPC}}_{it}$	0.071	0.121
Rival adoption: $\overline{\text{UPC}}_{it}$	0.161	0.170
Trademark: TM_{it}	0.016	0.117
Ever TM	0.081	0.266
Firms ^b	779,300	
Observations ^b	5,112,400	

Notes: ^aRevenue, reported in millions of 1992 dollars, is based on a different sample. See text for details.

^bAn observation is a firm-year except for the productivity sample. Firm and observation counts rounded to comply with Census rules on disclosure avoidance.

Table 2. UPC Diffusion Hazard Regressions

Spillover	Channel	Industry	Channel	Industry	ZIP
Channel: $\widehat{\text{UPC}}_{it}$	0.0276*** (0.0049)		0.0196*** (0.0045)		
Rival: $\overline{\text{UPC}}_{it}$		0.0261*** (0.0023)		0.0125*** (0.0015)	-0.0004 (0.0004)
SIC/ZIP fixed effects			✓	✓	✓
Value added			✓	✓	
Mean outcome			0.0034		
Observations ^a			5,033,100		

Notes: Outcome: UPC adoption. Firms remain in sample until year of first UPC adoption.

Robust SEs clustered by firm SIC (ZIP) in parentheses. * p<10%; ** p<5%; *** p<1%

All models control for firm-age×year effects, $\ln(\text{Employment}_{t-1})$, and a vertical-integration indicator.

^aAn observation is a firm-year. Observation counts rounded to comply with Census rules on disclosure avoidance.

Table 3. Difference-in-Difference Regressions:
Employment and Revenue

Outcome	Employment		Revenue	
	Full	Matched	Full	Matched
UPC_{it}	0.155*** (0.007)	0.130*** (0.014)	0.204*** (0.011)	0.103*** (0.022)
Observations ^a	5,112,400	221,600	1,113,000	44,000

Notes: Outcomes are logged. * p<10%; ** p<5%; *** p<1%. Robust SEs clustered by four-digit firm SIC in parentheses.

All regressions include firm, four-digit SIC×year, and firm-age×year fixed effects.

^aAn observation is a firm-year. Observation counts rounded to comply with Census rules on disclosure avoidance.

Table 4. Network Effect Difference-in-Difference Regressions

	Employment		Revenue	
UPC_{it}	0.092*** (0.022)	0.100*** (0.018)	0.045 (0.040)	0.073*** (0.027)
$UPC_{it} \cdot \widehat{UPC}_{it}$	0.294*** (0.087)		0.279** (0.138)	
$UPC_{it} \cdot \overline{UPC}_{it}$	0.175*** (0.063)		0.262** (0.118)	
Observations ^a	221,600	221,600	44,000	44,000

Notes: Outcomes are logged. Robust SEs clustered by four-digit firm SIC in parentheses. * p<10%; ** p<5%; *** p<1%.

All regressions include firm, firm-age×year, and industry×year fixed effects.

^aAn observation is a firm-year. Observation counts rounded to comply with Census rules on disclosure avoidance.

Table 5. Difference-in-Difference Regressions:
Trademarking

Sample	Full	Matched
UPC_{it}	0.044*** (0.003)	0.045*** (0.003)
Observations ^a	5,112,400	331,000

Notes: Outcome: indicator for firm trademarking. Robust SEs clustered by four-digit firm SIC in parentheses. * p<10%; ** p<5%; *** p<1%. All regressions include firm, firm-age×year, and industry×year fixed effects.

^aAn observation is a firm-year. Observation counts rounded to comply with Census rules on disclosure avoidance.

Table 6. Industry-Level Trade Regressions

	Imports		Import Penetration	
\widehat{UPC}_{mt}	0.7698*** (0.3227)		0.2766*** (0.0480)	
\overline{UPC}_{mt}		0.5204*** (0.1346)		0.0549*** (0.0143)
Controls ^a	✓	✓	✓	✓
Observations ^b	7,596			

Notes: Imports are logged. Robust SEs clustered by four-digit SIC in parentheses. * p<10%; ** p<5%; *** p<1%.

^aAll regressions include industry and year fixed effects and controls for log (domestic) industry value added, log capital-labor ratio, and log production-to-non-production worker ratio.

^bAn observation is an industry-year.

A Data Construction

A.1 Matching UPC Registrations to Census Establishments

The final datasets and source code for this project may be accessed, by request, by researchers on approved projects using confidential data in the Federal Statistical Research Data Center network whose projects include the underlying source datasets (BR/SSEL, LBD, CM, CW, CRT, and ASM).

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.⁴⁷ 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.⁴⁸ For each of the establishments, we use firm identifier in these files to identify any other 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

⁴⁷Exceptions include most government-owned or -operated establishments, establishments operated by religious organizations, and agricultural establishments.

⁴⁸Almost 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.

through 2000 from the Business Register (BR).⁴⁹

We match the UPC registrations to business names and addresses in the BR.⁵⁰ 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 physical address) each year. 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.

Figure A-1 shows the overall match rate by year and firm size. The match rate begins around 75 percent, when most UPC adopters were relatively large firms, before declining to roughly 40 percent in the 1980s when a large majority of registrants had less than \$2 million in annual revenues. We attempted to improve match rates by adding BR records from outside the manufacturing, wholesale and retail sectors of the economy. Specifically, we expanded our universe to firms in business services (SIC 73), legal services (SIC 81), and transportation and warehousing (SIC 40), but found that the number of additional

⁴⁹Most registrations are in the manufacturing and wholesale sectors. Some retailers registered for a UPC as well, for several reasons. Some retailers registered to show support for the system. Others owned upstream establishments, such as a production facility for private-label items. Grocery stores with meat and deli departments may have registered for a UPC in order to print their own barcodes on variable-weight items (Selmeier, 2008, pp. 131–132).

⁵⁰We match registrations from 1974 and prior years to the 1975 BR.

matches was very small. Nevertheless, our matching results are in line with similar efforts by other authors. For example, Brown and Earle (2013, Table 1) obtain a 44 percent match rate between employer records in the BR and a sample of firms that applied for loans from the Small Business Administration, consistent with the match rate for smaller firms in our sample. For larger firms, our match rates are similar to those reported by Jarmin (1999) for manufacturing plants or Kerr and Fu (2008) for patent filers.

A.2 Aggregation to Firm Level

Most firms operate just one establishment, and in those cases constructing firm-level variables is trivial. For multi-unit firms, we identify a new UPC registration in year t when a firm that did not have a registration in year $(t - 1)$ 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 each establishment to an industry 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.⁵¹ To assign a single, time-invariant SIC to the firm, we use the firm's employment, payroll (deflated by the CPI), and establishment count summed over its lifetime. Specifically, we first sum firm employment across all establishments and all years by sector; we assign the firm to the sector — manufacturing, wholesale, retail, or other — with the plurality of the firm's total employment. Next, we use total firm employment over the full time period to determine the *two-digit*

⁵¹The Census switched to the North American Industrial Classification System (NAICS) starting in 1997. Our sample ends in 1992, so we rely on SIC codes.

SIC within that sector. Within that two-digit SIC, we use the same procedure to assign a three-digit SIC, and then a four-digit SIC. In each step we break ties (multiple SICs with identical employment) using the firm’s CPI-deflated total payroll, and then establishment count, over its lifetime. This procedure ensures that, for example, a firm whose employment is predominantly in food manufacturing, but spread across multiple four-digit industries, but which has one large plant in apparel manufacturing, receives a food-manufacturing SIC code.⁵²

Employment at the firm level is available at annual frequency, by aggregating the LBD employment figures across all establishments the firm operates in that year. In Economic Census years we also have establishment-level revenue, which we aggregate to the firm level. Firm revenue includes revenue reported in the CM, CW, and CRT only.

A.3 $\overline{\text{UPC}}_{it}$ and $\widehat{\text{UPC}}_{it}$ Calculations

Recall that the definition of $\overline{\text{UPC}}_{it}$ is given by Equation (1):

$$\overline{\text{UPC}}_{it} = \sum_{e \in E(i)} w_e \left[\sum_{k \in M(e)} z_k \text{UPC}_{kt} \right] \quad (1)$$

where $E(i)$ denotes all establishments at firm i , and the set $M(e)$ contains all establishments in the same four-digit SIC code as e , excluding any establishments owned by firm i . The

⁵²We use a similar procedure to assign a firm-level ZIP code. First, we aggregate firm employment by ZIP code across all years the firm exists in the sample. Second, we find the one-digit ZIP code with the highest total employment for the firm. Third, within that one-digit ZIP code, we find the two-digit ZIP code with the highest total employment. Finally, within that two-digit ZIP code, we find the three-digit ZIP code with the highest employment. Here, too, we break ties using payroll and establishment counts.

employment weights are:

$$w_e = \frac{\text{Employment}_e}{\sum_{j \in E(i)} \text{Employment}_j} \quad (\text{A-1})$$

$$z_k = \frac{\text{Employment}_k}{\sum_{j \in M(e)} \text{Employment}_j} \quad (\text{A-2})$$

The weights used in the construction of $\widehat{\text{UPC}}_{it}$, and the process used to create them, are more complex. The definition of $\widehat{\text{UPC}}_{it}$ from Equation (2) is

$$\widehat{\text{UPC}}_{it} = \sum_{r \in R} s_{jr} \left[\sum_{m \in \{M \setminus j\}} u_{rm} \overline{\text{UPC}}_{mt} \right] \quad (2)$$

where R is the set of all four-digit retail SICs; $\{M \setminus j\}$ is the set of all manufacturing industries except for j (the industry containing firm i); and $\overline{\text{UPC}}_{mt}$ is the employment-weighted industry average UPC adoption for manufacturing industry m in year t . To create $\overline{\text{UPC}}_{mt}$ we follow the same steps used to construct $\widehat{\text{UPC}}_{it}$, but without excluding any establishments from the set $M(e)$.

To construct the weights s_{jr} and u_{rm} , we create a concordance between manufacturing SIC codes and broad merchandise-line codes from the CRT.⁵³ Each store in the CRT is asked to report its revenue by broad merchandise line. Examples of broad lines are food, women's apparel, and furniture. Basker, Klimek, and Van (2012) describe the lines data in more detail. For each broad merchandise line, we identify all upstream (manufacturing and wholesale) four-digit SIC codes supplying that line based on descriptions of the products manufactured by establishments in that SIC code. We also use data on revenue by merchandise line sold through each four-digit retail SIC, as reported in the 1977 CRT.

Let L denote the set of all merchandise-line codes. From the 1977 CRT tabulations we obtain LineRevenue_{rl} , the total revenue associated with line l in retail industry r . Using

⁵³Our concordance is available online at <http://people.bu.edu/tsimcoe/data/>.

these data, we calculate the share of each line l in industry r :

$$\text{Share}_{rl} = \frac{\text{LineRevenue}_{rl}}{\sum_{j \in L} \text{LineRevenue}_{rj}}.$$

Now define $L(m)$ to be the set of all merchandise lines supplied by manufacturing industry m according to our concordance, and N_l to be the total number of manufacturing industries supplying line l . Using these expressions, the inner weights in the expression for $\widehat{\text{UPC}}_{it}$ are

$$u_{rm} = \sum_{l \in L(m)} \frac{\text{Share}_{rl}}{N_l}. \quad (\text{A-3})$$

These weights provide an approximation of the share of retail industry r 's revenue attributed to lines produced by manufacturing industry m .

The outer weights s_{jr} represent the share of revenue that manufacturing industry j (the industry containing firm i) derives from retail industry r . Once again, we rely on the CRT-SIC concordance to produce this weight. As a first step, we use the concordance to calculate the total revenue collected by manufacturing industry j from retail channel r

$$\text{Revenue}_{jr} = \sum_{l \in L(j)} \frac{\text{LineRevenue}_{rl}}{N_l}$$

And in a second step, we construct the outer weights by summing over all downstream retail channels:

$$s_{jr} = \frac{\text{Revenue}_{jr}}{\sum_{k \in R} \text{Revenue}_{jk}} \quad (\text{A-4})$$

Illustrative Example

To illustrate how we calculate $\widehat{\text{UPC}}_{it}$, consider the following hypothetical example, based on an economy with three manufacturing industries (food, apparel, and other) and two retail industries (grocery and department store). For simplicity, we begin by assuming a one-to-one correspondence between manufacturing industries and broad lines, so $u_{rm} \equiv \text{Share}_{rl}$ and

$\text{Revenue}_{jr} \equiv \text{LineRevenue}_{rl}$. We consider a more complex case below. The following table illustrates the raw data used in our computations, where the first two columns would be obtained from the 1977 CRT, and the third column would be computed from our panel of UPC registrations.

	Grocery	Department	$\overline{\text{UPC}}_{mt}$
Food	\$90M	\$20M	60%
Apparel	\$5M	\$50M	30%
Other	\$5M	\$30M	10%

To calculate $\widehat{\text{UPC}}_{it}$ for apparel manufacturing, we can apply Equation (2) as follows:

$$\widehat{\text{UPC}}_{\text{apparel},t} = \frac{5}{55} \left[\frac{90}{95} * 0.6 + \frac{5}{95} * 0.1 \right] + \frac{50}{55} \left[\frac{20}{50} * 0.6 + \frac{30}{50} * 0.1 \right] = 0.325 \quad (\text{A-5})$$

Note that the outer weights, s_{jr} , correspond to the share of total apparel sales through each retail channel, whereas the inner weights, u_{rm} , correspond to each industry's share of total revenue (excluding apparel) within that channel.

The many-to-many correspondence between broad lines and manufacturing SIC codes creates two types of complications: a single line matching to many SIC codes, or a single SIC code matching to many lines. In the latter case, we simply aggregate over broad lines, using either Equation (A-3) or the definition of Revenue_{jr} , both of which apportion revenue equally across all SIC codes that supply a particular broad line.

To see how we handle the case where a broad line corresponds to multiple SICs, suppose that the rows in the example table, which so far represented both retail lines and manufacturing industries, now represent only retail lines. The food line is supplied by two industries: canned-foods producers and fresh-foods producers, with UPC adoption rates of 80% and 20%, respectively. In that case, Equation (A-3) implies that the first bracket term in Equation (A-5) becomes

$$\left[\frac{1}{2} * \frac{90}{95} * 0.8 + \frac{1}{2} * \frac{90}{95} * 0.2 + \frac{5}{95} * 0.1 \right] = 0.044$$

and making the same adjustment to the second term in brackets yields $\widehat{\text{UPC}}_{\text{apparel},t} = 0.280$.

A.4 Aggregate TM Registrations in the Grocery Sector

Figure 6 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 restrict the sample to grocery-related TM categories and account for missing data prior to 1977.⁵⁴

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 one or more three-digit Nice code corresponding to food, beverages, pharmaceuticals or paper products.

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 denotes grocery trademarks applications in year t , we create a new variable:

$$\widehat{\text{TM}}_t = \text{TM}_t \times \frac{\text{TM}_{1977}}{\text{TM}_{1974}}$$

⁵⁴Data and code for replicating Figure 6 are available at <http://people.bu.edu/tsimcoe/data.html>.

and replace TM_t with \widehat{TM}_t for all years prior to 1975. Second, we linearly interpolate values for 1975 and 1976 – the filing years most influenced by the change in USPTO record keeping – and plot the resulting time-series in Figure 6.

Neither of these adjustments is made to the data used in our firm-level TM analysis in Section 5.2 and Appendix A.4, which includes both grocery- and non-grocery firms and uses calendar-year fixed effects to control for the change in PTO procedures.

B Supplemental Empirical Results

B.1 Scanner Diffusion

Although we have only limited data on scanner adoption, this appendix provides some additional evidence of network effects on the retail side of the UPC platform by estimating models of grocery-store scanner adoption using data from Basker (2012). We conduct the analysis at the store (i.e., establishment) level rather than chain (firm) level because we have store-level data, and because scanner adoption occurred gradually within large retail chains.⁵⁵

For this analysis, we create a store-specific measure of upstream UPC adoption within a retail supply chain. Because we do not observe actual buyer-supplier relationships, our measure combines industry-level variation in UPC adoption with store-level data on the merchandise mix offered by individual retailers. Specifically, we use the CRT-SIC concordance between upstream industries and broad merchandise lines described in Appendix A.3 to construct a variable, $Upstream_{st}$, that captures the revenue-weighted average UPC adoption rate of industries supplying store s at time t .

To create our store-level measure of upstream UPC adoption, we start with the industry-level adoption rate \overline{UPC}_{mt} for each four-digit manufacturing industry m . Using our corre-

⁵⁵Basker (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.

spondence between SIC codes and broad merchandise lines, we then compute an employment-weighted supplier adoption rate UPC'_{lt} for each merchandise line l . The explicit formula is

$$UPC'_{lt} = \frac{\sum_{k \in M(l)} \overline{UPC}_{kt} \text{Employment}_{kt}}{\sum_{k \in M(l)} \text{Employment}_{kt}},$$

where $M(l)$ is the set of four-digit industries supplying merchandise line l , and Employment_{kt} is total employment at all establishments in manufacturing industry k in year t . Finally, for each store s , we compute a revenue-weighted UPC adoption rate across all lines $l(s)$ sold by that establishment.⁵⁶

$$\text{Upstream}_{st} = \sum_{k \in L(s)} \frac{\text{Revenue}_{kt} UPC'_{kt}}{\text{Revenue}_{st}} \quad (\text{B-1})$$

Table B-1 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 install a scanner by 1984. The latter sample allows us to isolate the effects of upstream UPC adoption on the timing of scanner installation, at the cost of selecting on the outcome variable. Both panels contain store-year observations up to and including the year of scanner installation, at which time a store is removed from the sample.

The first row of Table B-1 shows that the mean hazard of grocery-store scanner adoption between 1974 and 1984 was 0.8 percent. Overall, 3.6 percent of the grocery store-year observations belong to an establishment that 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 during our sample period.

In a simple model of adoption, firms compare the costs of installing scanners in a store

⁵⁶In inter-censal years, we use the stores' revenue shares from the prior CRT.

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:

$$\text{Scanner}_{st} = \lambda_{at} + \beta \text{Upstream}_{st} + X_{st}\theta + \varepsilon_{st} \quad (\text{B-2})$$

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

For each of the two samples, we first report a minimalist regression in which we control only for store age by calendar year fixed effects, and then a regression that controls for employment at both the store and the chain to which it belongs; an indicator for vertical integration (i.e., the store is part of a firm that owns at least one wholesale or manufacturing plant); and an indicator for the owning firm’s UPC registration. Because stores that have installed scanners drop out of the sample in subsequent years, β corresponds to a change in the hazard of scanner adoption. For all models, we cluster standard errors at the store level.

Table B-2 presents our scanner-diffusion estimates. The coefficient estimate $\hat{\beta}$ is positive and statistically significant across all samples and specifications, consistent with the presence of indirect network effects. To interpret the magnitude of this coefficient, note that a one-standard-deviation increase in Upstream_{st} is associated with a 23 percent increase in the baseline hazard (i.e., the sample mean adoption rate reported at the bottom of Table B-2). The unreported coefficients on firm-level controls indicate higher scanner adoption rates

for larger firms, consistent with the presence of chain-level economies of scale in scanner deployment.

Estimates of β 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 3.2 percent. Thus, although our measure of upstream UPC adoption is positively correlated with scanner adoption in both samples, it explains more about which stores installed scanners (before 1984) than about how quickly they did so.

B.2 UPC Adoption and Survival

There are several reasons why UPC adoption might be correlated with an increased likelihood of firm survival. The employment and revenue models in Section 5 suggest that UPC adoption occurs when firms find new markets, which itself should enhance the odds of survival. UPC adoption may also indicate growth or technology adoption within a manufacturer’s distribution channel. Moreover, firms anticipating exit are unlikely to invest in new technologies, so future survival may predict UPC adoption. In this section, for completeness, we provide results from a series of exit regressions.

Our primary outcome variable is a binary variable Exit_{it} , which equals one for the last year any firm appears in the sample. (We drop data for 1992 because we cannot observe whether a firm is in the data in 1993). Table 1 shows that the mean hazard for the full sample is 9.7 percent. We estimate the following specification, which is similar to the adoption hazard models:

$$\text{Exit}_{it} = \lambda_{at} + \beta \text{UPC}_{it} + \varepsilon_{it} \tag{B-3}$$

where λ_{at} are firm-age by year fixed effects, and standard errors are clustered at the firm-level.

Results for the full sample appear in the first two columns of Table B-3. The coefficient β can be interpreted as the change in the hazard of exit associated with UPC adoption.

Not surprisingly, this coefficient is negative and significant. The estimate in the first column implies a 2.3 percentage point difference in the exit hazard for a firm with a UPC registration compared to one without a UPC registration — a decline of approximately 25% in the baseline hazard rate. In the second column, we add manufacturing industry by year fixed effects to allow for differential exit trends by four-digit firm SIC. This produces a small increase in the magnitude of the coefficient on UPC adoption.

As we discuss in Section 5.1, the control group of non-adopters in the full sample differs from UPC adopters on a number of dimensions, including size and prior growth. To examine the relationship between UPC adoption and exit in the matched sample, we re-estimate the hazard models using the employment-matched sample. For this analysis, we drop all years prior to UPC adoption, since both the adopter and its matched control are guaranteed to survive up to that point in time. The coefficients from these regressions, both without and with industry-year fixed effects, are in the last two columns of Table B-3. Although the association between UPC adoption and firm exit is somewhat smaller than in the full sample, the coefficients in the third and fourth columns of Table B-3 remain economically and statistically significant. Overall, we conclude that UPC adoption is associated with a meaningful increase in the probability of firm survival relative to similarly sized non-adopters (in different industries) that grew at a similar rate prior to adoption.

B.3 UPC Adoption and Productivity

The difference-in-difference estimates reported in Table 3 indicate that UPC adoption was followed by an increase in both revenue and employment, and that these two effects were similar in magnitude. These findings suggest that for manufacturers, UPC adoption had at most a modest impact on productivity. The absence of any large upstream productivity effect is consistent with our reading of the historical record. The UPC was pushed by retailers with the goal of increasing productivity at checkout, and was adopted with some reluctance by manufacturers, which saw it as a net burden.

To test formally for a relationship between UPC adoption and productivity, we use a sample of manufacturing establishments, rather than firms, because manufacturing productivity is traditionally estimated at the plant level. Our sample consists of all manufacturing establishments in the Census of Manufactures (1977, 1982, 1987, and 1992) supplemented by observations from the Annual Survey of Manufactures (ASM) for inter-censal years. The ASM is a rotating panel of roughly 50,000–70,000 establishments. Establishments remain in the ASM for five years, with some overlap across panels, particularly for establishments with at least 250 employees.⁵⁷

We take the Total Factor Productivity (TFP) measure computed by Foster, Grim, and Haltiwanger (2016) as our outcome variable, and estimate a two-way fixed effects specification

$$\text{TFP}_{eit} = \alpha_e + \delta_t + \beta \text{UPC}_{it} + \varepsilon_{et} \quad (\text{B-4})$$

where α_e is an establishment fixed effect, δ_t is a year fixed effect, and UPC_{it} is the UPC status of the firm to which establishment e belongs. The TFP measure is in logs, and already accounts for changes over time in capital and labor inputs.⁵⁸ As additional time-varying controls, we estimate models that include establishment-age by year fixed effects, and industry by year fixed effects. For each regression, we report two sets of standard errors: first unclustered and then clustered at the four-digit (establishment) SIC level. Because we are interested in the null hypothesis of no relationship between UPC adoption and TFP, the unclustered standard errors are conservative — they provide the greatest chance of a rejection.

TFP results are reported in Table B-4. The first column has establishment and year fixed effects. In the second column we add firm-age by year fixed effects and the third column

⁵⁷The ASM is not a representative sample: establishments are selected based on industry and size. Our estimates are all unweighted.

⁵⁸By construction, TFP has mean zero. The mean of exponentiated TFP in the sample is 1.820, with standard deviation 0.599.

adds industry by year effects. The difference-in-difference estimates of the impact of UPC adoption on TFP range from -0.15 percent to 0.23 percent across our three regressions. None of these results is statistically significant at conventional levels, even if we focus on the unclustered standard errors. We do not use a matched sample with this specification because the selection concern is that firms with higher-productivity establishments may register for a UPC earlier or at a higher rate than firms with low-productivity establishments. If that is the case, our reported coefficient estimates are biased upwards.

Overall, the TFP results in Table B-4 suggest that the historical conventional wisdom was correct. Whatever benefits retailers realized from barcodes, UPC adoption did not produce large changes in manufacturing productivity. This interpretation comes with one major caveat — we use revenue rather than quantity TFP. It is possible, therefore, that manufacturing plants did become more productive, but also gave price concessions to retailers, so that changes in revenue understate the total impact of UPC adoption on output quantities. We lack the quantity data that would be required to properly test this idea. But even if it were true, that finding would simply reinforce the idea that the major benefits of the UPC system (and also the greater investments in IT and organization change) occurred downstream in the retail sector.

C Replication for Wholesalers

Many early UPC adopters were outside the manufacturing sector. In this appendix, we replicate our main results using a sample of firms in the wholesale sector. 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 a large majority of wholesaler UPC

registrations belong to the merchant category.⁵⁹ Dinlersoz, Goldschlag, Myers, and Zolas (forthcoming) show that wholesalers are among the firms most likely to pursue trademarks, and Ganapati (2016) studies the role of merchant wholesalers in the supply chain.

The data and variable construction for this wholesaler analysis largely parallels the steps taken for manufacturers, and readers should consult the main text and Appendix A for details. There are two notable differences. First, because the LBD starts in 1975, firm age is censored at $(t - 1975)$ for wholesalers. We are able to go back to 1972 only for manufacturers that appeared in the 1972 Census of Manufactures. Second, the SIC codes are coarser for the wholesale sector than for manufacturing. The 1988 SIC file had 577 four-digit manufacturing codes, compared to 87 four-digit wholesale codes. This means that we have a smaller number of clusters in many analyses, and there is less between-industry variation in our key measures of adoption, $\widehat{\text{UPC}}_{it}$ and $\overline{\text{UPC}}_{it}$.

Table C-1 provides summary statistics for the wholesaler sample. Wholesale firms are significantly smaller than manufacturing firms, averaging just 15 employees, compared to 73 in the manufacturer sample. UPC adoption rates are around half of those observed in manufacturing sample: 2.2% of the observations in our wholesale sample correspond to a firm that registered for a UPC by 1992, compared to 3.8% in the manufacturer sample. For wholesalers, the average rival UPC adoption, $\overline{\text{UPC}}_{it}$, is a bit higher than for manufacturers, whereas channel adoption, $\widehat{\text{UPC}}_{it}$, is a bit lower. Finally, Table C-1 shows that wholesalers file new trademark applications at around half the rate of manufacturers, and that firms in the two samples have a very similar exit rate.

Figure C-1 shows the diffusion of the UPC in the wholesale sector, across years and firm-size quartiles. The general pattern is the same one that we observed in the manufacturing sector: firms in the top quartile of the firm size distribution are more likely to register for a

⁵⁹Starting in 2002, NAICS codes have distinguished between merchant wholesalers and branch offices. That year, merchant wholesalers accounted for 93 percent of wholesale establishments and 90 percent of wholesale revenue (U.S. Census Bureau, 2005).

UPC, with an increasing gap in cumulative registration rates over time. Consistent with the summary statistics, adoption rates within each quartile are around half of the manufacturing registration rates displayed in Figure 2. (Firms in the retail sector have lower registration rates, but display a similar pattern of UPC adoption by firm-size quartile.)

Table C-2 shows estimates from hazard models, based on Equation (6), using the wholesaler sample. As in the manufacturer analysis, all of these regressions include firm-age by year fixed effects, lagged employment, and a vertical-integration indicator as controls, and standard errors are clustered at the four-digit firm SIC level.⁶⁰ The baseline hazard of UPC registration among wholesalers during our sample period was 0.23 percent. The first two columns in Table C-2 are based on a pure correlated-effects specification that does not include industry fixed effects. We find a positive and statistically significant coefficient in both models, consistent with the presence of network effects. A one-standard-deviation change in the adoption variable is associated with an 86 percent increase in the baseline hazard of UPC registration for rival adoption, and a 59 percent increase in the hazard for channel adoption. These effects are somewhat smaller than the ones we estimated for the manufacturer sample. The third and fourth columns of Table C-2 show estimates from a specification that includes SIC fixed effects. Relative to the first two columns, the estimate for rival adoption falls by around 25 percent, and the estimate for channel adoption declines by around 40 percent. Although the coefficient on $\widehat{\text{UPC}}_{it}$ in the latter model loses statistical significance, it still implies that a one-standard-deviation change in channel adoption increases the hazard of UPC registration by 32 percent.

Table C-3 reports coefficient estimates from difference-in-difference models, using log firm employment as the outcome variable. The results in the first two columns are based on the specification of Equation (7). The full sample results show a 20 percent increase in employment following UPC registration, which is somewhat larger than the manufacturing

⁶⁰For the wholesaler sample, the vertical-integration indicator turns on if the firm has at least one retail or manufacturing establishment.

estimates reported in Table 3. In the second column, we report estimates for a matched sample constructed according to the same procedure we use for manufacturers. Again, the 17 percent increase is a bit larger than the 13 percent increase we find in the matched sample of manufacturing firms. One explanation for finding larger effects among wholesalers is that they are smaller on average, so hiring one or two employees is a proportionately larger change.

The third and fourth columns in Table C-3 show results from a network-effects specification, based on Equation (9). In these models, our results diverge from the manufacturing estimates reported in Table 4. Neither of the two interactions (with channel and rival adoption, respectively) are statistically significant. With lower overall UPC adoption rates in the wholesale sector, it is possible that network effects were less important for firm-level outcomes (although the estimates in Table C-2 suggest that they did matter for diffusion). Another possibility is that there is more measurement error in both \widehat{UPC}_{it} and \overline{UPC}_{it} , given the less granular SIC codes in wholesaling, making it difficult to estimate a precise effect.

Figure C-2(a) shows event-study coefficients, based on Equation (8), for the matched samples of wholesalers. Here, the pattern is quite similar to what we find in manufacturing: there is a sharp increase in employment in the year of UPC registration, followed by a more gradual increase over the following years. In Figure C-2(b) we show the event-study coefficients for the full sample. These estimates show strong evidence of selection effects: UPC adopters grow more quickly than non-adopters in the years leading up to their first UPC registration.

Finally, Table C-4 reports coefficients from difference-in-difference regressions using an indicator for new TMs as the outcome. The results are very similar to those for manufacturing. In the full sample, UPC registration is associated with a 3.5 percentage point increase in TM filings, and the corresponding matched sample estimate suggests a 4.5 percentage point increase. To provide a sense of economic magnitudes, we can take the ratio of TM_{it} and Ever Trademark in Table C-1 to infer that the annual probability of filing a new TM

application among firms that ever do so is around 17 percent. Thus, a 4.5 percentage point increase in the filing rate corresponds to a marginal effect of 26 percent.

Overall, the estimates for wholesalers across all of the models we estimate are very similar to the results for manufacturers that we report in the paper. The main exception are the statistically insignificant interactions terms in our network effects specification (i.e., the third and fourth columns Table C-3). Nevertheless, we view the evidence from this replication of our manufacturing analysis on a sample of firms in the wholesale sector as largely confirming our main firm-level findings regarding UPC adoption and its impacts.

D Derivation of Diffusion Models

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, let S be the population of stores and s_t the number that have installed scanners. We assume 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:

$$\dot{m}_t = \alpha_m s_t \tag{D-1}$$

$$\dot{s}_t = \alpha_s m_t \tag{D-2}$$

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

$$m_t = \exp\{(\alpha_m \alpha_s)^{\frac{1}{2}} t\} \quad (\text{D-3})$$

$$s_t = \left(\frac{\alpha_s}{\alpha_m}\right)^{\frac{1}{2}} \exp\{(\alpha_m \alpha_s)^{\frac{1}{2}} t\} \quad (\text{D-4})$$

Substituting Equations (D-3) and (D-4) into Equation (D-1) yields the following reduced-form specification for UPC adoption in levels:

$$\dot{m}_t = (\alpha_m \alpha_s)^{\frac{1}{2}} m_t \quad (\text{D-5})$$

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:

$$h(m_t) = \frac{\dot{m}_t}{M - m_t} = (\alpha_m \alpha_s)^{\frac{1}{2}} \frac{m_t}{M - m_t} = (\alpha_m \alpha_s)^{\frac{1}{2}} \frac{\overline{\text{UPC}}}{1 - \overline{\text{UPC}}} \quad (\text{D-6})$$

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

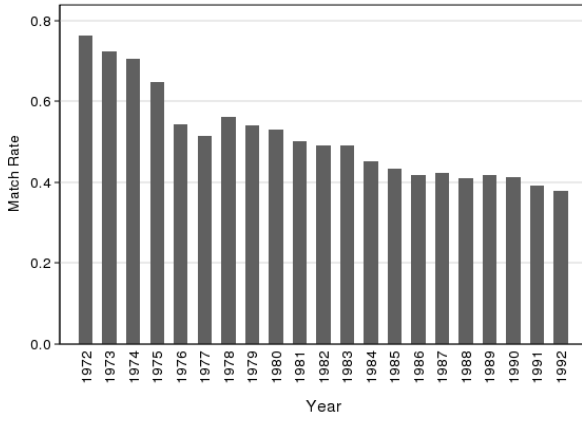
Equation (D-6) is nearly identical to the reduced-form specification in Equation (6), except that we replace the nonlinear function $\frac{\overline{\text{UPC}}}{1 - \overline{\text{UPC}}}$ with $\overline{\text{UPC}}$. This makes our specification equivalent to the well-known Bass (1969) diffusion model, and more importantly, simplifies the interpretation of the coefficient on $\overline{\text{UPC}}$. We have estimated versions of Equation (6) 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.

It is much simpler to derive the scanner-diffusion specification, Equation (B-2) used in Appendix B.1, because we have a direct measure of complementary UPC registrations: Upstream_{st} . Dividing both sides of (D-2) by $(S - s_t)$ yields

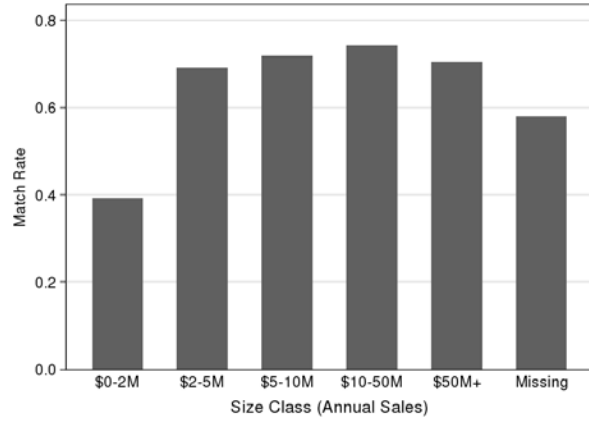
$$h(s_t) = \frac{\dot{s}_t}{S - s_t} = \alpha_s \cdot \frac{m_t}{S - s_t} \approx \frac{\alpha_s}{S} \cdot \text{Upstream}_t \quad (\text{D-7})$$

where the last step relies on the fact that scanner adoption was negligible, so $s_t \approx 0$, even among food stores. Specifically, Table B-1 indicates that just 3.7 percent of the observations in our food-store panel are from stores that installed scanners by 1984.

Figure A-1. UPC-BR Match Rate



(a) By Year



(b) By Size Class

Figure C-1. UPC Diffusion by Firm Revenue, Wholesale

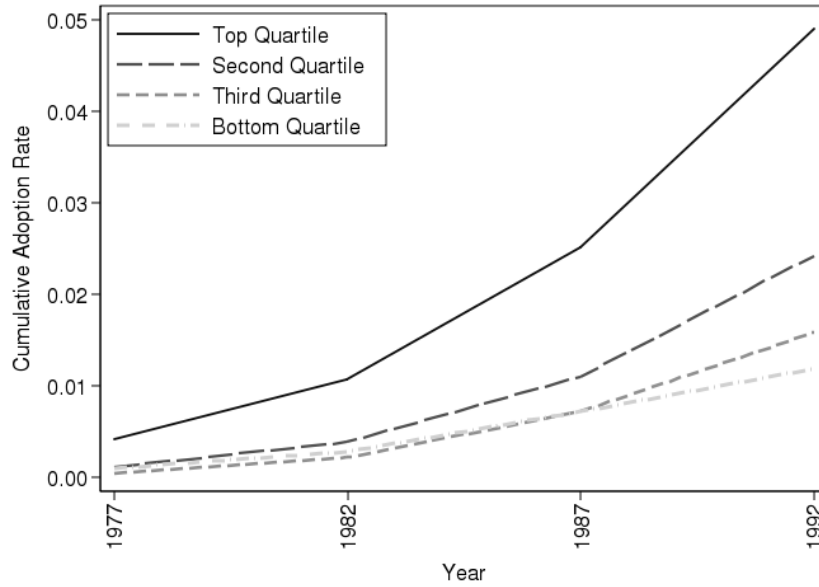
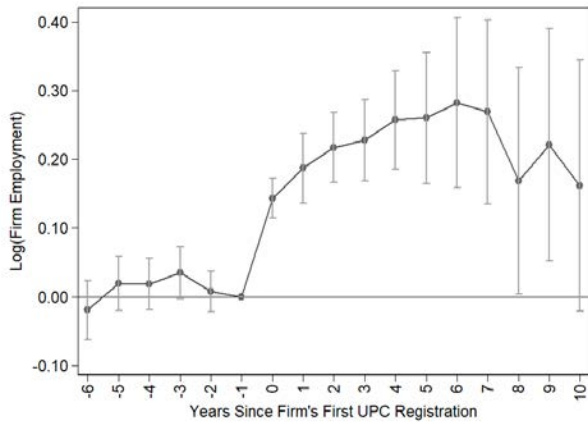
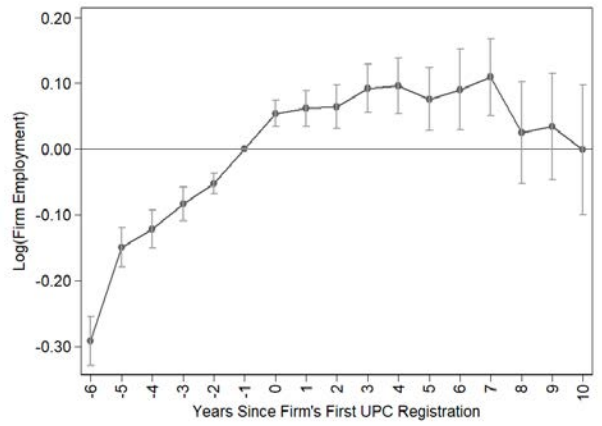


Figure C-2. Event Study Coefficients: Employment, Wholesale



(a) Matched Sample



(b) Full Sample

Note: Vertical bars represent 95 percent confidence intervals

Table B-1. Summary Statistics, Retail

	Food Stores		Scanning Stores	
	Mean	SD	Mean	SD
Scanner _{st}	0.008	0.087	0.213	0.410
EverScans ^a	0.036	0.187	1.000	0.000
Upstream _{st}	0.356	0.070	0.376	0.027
Stores ^b	89,500		3,300	
Observations ^b	418,000		15,000	

Notes: Scanning stores installed a front-end scanner by 1984 (Basker, 2012). Food stores include all scanning stores and other stores in SIC 5411.

^aEverScans is equal to 1 if the store installed a scanner by 1984.

^bAn observation is a store-year. Store and observation counts rounded to comply with Census rules on disclosure avoidance. Stores remain in the sample until year of scanner adoption or 1984, whichever is earlier.

Table B-2. Scanner Adoption Hazard Regressions

	Food Stores		Scanning Stores	
Upstream _{st}	0.0632*** (0.0015)	0.0260*** (0.0013)	0.2864** (0.1380)	0.2551* (0.1409)
Controls ^a		✓		✓
Mean outcome	0.0078		0.2126	
Observations ^b	418,000		15,000	

Notes: Outcome: Scanner adoption. Scanning stores adopt scanners by 1984. Stores remain in the sample until year of scanner adoption or 1984, whichever is earlier. 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.

^aControls include log store employment, log firm employment, a vertical-integration indicator, and an own-UPC registration indicator.

^bAn observation is a store-year. Observation counts rounded to comply with Census rules on disclosure avoidance.

Table B-3. Exit Hazard Regressions

	Full Sample		Matched Sample	
UPC _{it}	-0.023*** (0.003)	-0.029*** (0.002)	-0.019*** (0.004)	-0.020*** (0.003)
Industry×Year fixed effects		✓		✓
Observations ^a	4,800,000		102,000	

Notes: Robust SEs clustered by four-digit firm SIC in parentheses. * p<10%; ** p<5%; *** p<1%.

All regressions include firm age×year fixed effects. Matched regressions use observations starting with the (actual or counterfactual) year of UPC adoption.

^aAn observation is a firm-year. Observation counts rounded to comply with Census rules on disclosure avoidance.

Table B-4. Difference-in-Difference Regressions:
Total Factor Productivity

	Total Factor Productivity		
UPC _{it}	-0.0015 (0.0016) [0.0044]	-0.0003 (0.0017) [0.0040]	0.0023 (0.0019) [0.0038]
Establishment-age×year fixed effects		✓	✓
Industry×year fixed effects			✓
Observations ^a	1,292,000		

Notes: TFP measured in logs, with exponentiated mean 1.820 and standard deviation 0.599.

* p<10%; ** p<5%; *** p<1%. SEs: Unclustered in parentheses and clustered by industry in brackets.

All models contain establishment and year fixed effects.

^a An observations is an establishment-year. Observation counts rounded to comply with Census rules on disclosure avoidance.

Table C-1. Summary Statistics, Wholesale

	Mean	SD
Employees	14.9	193
UPC adoption: UPC_{it}	0.010	0.100
UPC adopter	0.022	0.145
Channel adoption: \widehat{UPC}_{it}	0.150	0.150
Rival adoption: \overline{UPC}_{it}	0.051	0.073
Trademark: TM_{it}	0.007	0.081
Ever TM	0.041	0.195
$\mathbb{I}[\text{Exit} \text{Alive}_{t-1}]$	0.096	0.295
Firms ^a	866,500	
Observations ^a	5,621,800	

Notes: ^aAn observation is a firm-year. Firm and observation counts rounded to comply with Census rules on disclosure avoidance.

Table C-2. UPC Diffusion Hazard Regressions, Wholesale

Spillover	Industry	Channel	Industry	Channel
Channel: \widehat{UPC}_{it}	0.0091*** (0.0018)		0.0049 (0.0045)	
Rival: \overline{UPC}_{it}		0.0268*** (0.0064)		0.0207*** (0.0048)
SIC fixed effects			✓	✓
Mean outcome		0.0023		
Observations ^a		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 in parentheses. * p<10%;

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

All regressions include firm-age×year fixed effects, $\ln(\text{Employment}_{t-1})$, and an indicator for vertical integration..

^aAn observation is a firm-year. Observation counts rounded to comply with Census rules on disclosure avoidance.

Table C-3. Difference-in-Difference Regressions:
Employment, Wholesale

Specification	Baseline		Network Effects	
	Full	Matched	Matched	Matched
UPC_{it}	0.217*** (0.013)	0.173*** (0.011)	0.197*** (0.020)	0.233*** (0.016)
$UPC_{it} \cdot \widehat{UPC}_{it}$			-0.057 (0.071)	
$UPC_{it} \cdot \overline{UPC}_{it}$				0.091 (0.071)
Observations ^a	5,621,800	166,500	166,500	166,500

Notes: Robust SEs clustered by four-digit firm SIC in parentheses.

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

\overline{UPC}_{it} is the adoption rate at the firm's competitors, by SIC

All regressions include firm, firm-age×year, and industry×year fixed effects.

^aAn observation is a firm-year. Observation counts rounded to comply with Census rules on disclosure avoidance.

Table C-4. Difference-in-Difference Regressions:
Trademarking, Wholesale

Sample	Full	Matched
UPC_{it}	0.035*** (0.003)	0.045*** (0.003)
Observations ^a	5,621,800	204,000

Notes: Outcome: indicator for firm trademarking.

Robust SEs clustered by four-digit firm SIC in parentheses. * p<10%; ** p<5%; *** p<1%.

All regressions include firm, firm-age×year, and industry×year fixed effects.

^aAn observation is a firm-year. Observation counts rounded to comply with Census rules on disclosure avoidance.

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, J., and J. S. Earle (2013) “Do SBA Loans Create Jobs?,” *IZA Discussion Paper No. 7544*.
- 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 (forthcoming) “An Anatomy of U.S. Firms Seeking Trademark Registration,” in *Measuring and Accounting for Innovation in the 21st Century*, ed. by C. Corrado, J. Miranda, J. Haskel, and D. Sichel.
- 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., C. Grim, and J. Haltiwanger (2016) “Reallocation in the Great Recession: Cleansing or Not?,” in *Labor Markets in the Aftermath of the Great Recession*, ed. by D. Card, and A. Mas, pp. 293–331. *Journal of Labor Economics*, Volume 34, Number S1, part 2.

- 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.
- GREENSTEIN, S., AND V. STANGO (eds.) (2007) *Standards and Public Policy*. Cambridge University Press.
- 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.
- Hillberry, R., and D. Hummels (2008) “Trade Responses to Geographic Frictions: A Decomposition Using Micro-Data,” *European Economic Review*, 52(3), 527–550.
- 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.
- Hsieh, C.-T., and E. Rossi-Hansberg (2019) “The Industrial Revolution in Services,” NBER Working Paper 25968.
- 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. (1999) “Evaluating the Impact of Manufacturing Extension on Productivity Growth,” *Journal of Policy Analysis and Management*, 18(1), 99–119.

- 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.
- Kerr, W. R., and S. Fu (2008) “The Survey of Industrial R&D–Patent Database Link Project,” *Journal of Technology Transfer*, 33(2), 173–186.
- 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.
- Nickell, S. (1981) “Biases in Dynamic Models with Fixed Effects,” *Econometrica*, 49(6), 1417–1426.
- 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.
- Rysman, M. (2019) “The Reflection Problem in Network Effect Estimation,” *Journal of Economics and Management Strategy*, 28(1), 153–158.
- 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.