european school
of management
and technology

September 27, 2007

ESMT
WORKING PAPER

ESMT No. 07-002

Estimating Level Effects In Diffusion Of A New Technology: Barcode Scanning At The Checkout Counter

Jonathan Beck, Humboldt Universität zu Berlin
Michał Grajek, ESMT
Christian Wey, Technische Universität Berlin

ISSN 1866-3494

Abstract

Estimating Level Effects in Diffusion of a New Technology: Barcode Scanning at the Checkout Counter

Author(s)*: Jonathan Beck, Humboldt Universität zu Berlin
Michał Grajek, ESMT
Christian Wey, Technische Universität Berlin

Cross-country or cross-industry studies of technology diffusion typically estimate how independent factors affect diffusion speed or timing, often based on a two-stage approach. In many applications, however, countries (industries) differ most in the saturation level of diffusion. In a novel, single-stage econometric approach to a standard diffusion model, we therefore estimate how the saturation level covaries with independent factors. In our application to diffusion of an important retail information technology, we focus on the competitive effect of hypermarkets (superstores). We also find standard scale, income and labor substitution effects.

Keywords: diffusion; information technology; retail competition

JEL Classification: L5, L81, O33

[1] Corresponding author: Jonathan Beck (beck@wz-berlin.de; postal address: WZB, Reichpietschufer 50, D-10785 Berlin).

Thanks to Joe Clougherty, Lapo Filistrucchi, Oz Shy, Irina Suleymanova and participants at the EEA conference in Vienna, the EARIE conference in Porto, the 5th ZEW Conference on the Economics of ICT in Mannheim, the BDPEMS workshop, the InterVal yearly meeting and the WZB seminar for comments and helpful discussions. Anna K¨alberer and Kemal Azun provided able research assistance. Financial support from the German Federal Ministry of Education and Research, project InterVal – Internet and value chains (01AK702A), is gratefully acknowledged. The authors are responsible for all remaining errors.

1 Introduction

Following the seminal work by Griliches (1957), cross-country or cross-industry studies of technology diffusion typically assess how diffusion speed or timing of a particular technology covaries with independent factors. The corresponding assumption is that the ceiling or saturation level of technology diffusion, which measures long-run technology usage, is invariant across time. While this assumption seems innocuous for technologies heading towards 100% penetration, such as Griliches' hybrid corn or, more recently, mobile telephony (Gruber and Verboven, 2001), it can be restrictive in other applications. In fact, Griliches (1957, p. 520) notes that "ceilings are not necessarily constant over time", and in a reply to late comments on his work he further states that he "would now use a model with an endogenous and shifting ceiling parameter", which the state of econometric art had prohibited earlier (Griliches, 1980, p. 1463).

To our knowledge, however, no study of technology diffusion has so far implemented this approach – despite the fact that cross-country level differences are considerable for many technologies (Comin, Hobijn and Rovito, 2006; Caselli and Coleman, 2001) and although, from a welfare perspective, long-run technology usage predominates diffusion speed.[1] The application presented in this paper is an exemplary case for endogenous saturation levels. In particular, we analyze newly compiled cross-country data on the diffusion of checkout barcode scanners in retailing, an information technology (IT) important also for other industries. The countries under consideration differ most in the long-run level of IT diffusion, less in timing or speed. We thus follow Griliches' recommendation to modify the standard diffusion model and estimate how independent factors affect the saturation level of diffusion. Particular focus is on the role of retail competition, namely in the form of hypermarkets.

In addition to a methodological point, this paper therefore has an interesting empirical contribution. In the context of our application, recent productivity studies attribute large post-1995 productivity gains in the United States to increased IT usage mainly in the distribution sector (Ark *et al.*, 2005; Stiroh, 2002). Most European countries, however, have not experienced such manifest developments in retail IT diffusion or productivity. An often-stated worry is that this may be due to excessive retail regulation and conse-

[1] Whereas Comin, Hobijn and Rovito (2006) study direct data on a large number of technologies and countries, Caselli and Coleman (2001) study the diffusion of computers using imports of computing equipment as an indirect measure.

quently less intensive retail competition in Europe (Scarpetta *et al.*, 2002, for example). Empirical results in this respect are of interest not only to understand historical developments, but also because the next generation of retail IT – radio frequency identification (RFID) – is on the verge of mass market introduction. Yet, empirical studies of the relationship between retail regulation and competition on the one hand and retail innovation and productivity on the other hand are rare, despite a rich parallel literature on the link between retail regulation and employment (Bertrand and Kramarz, 2002). More comprehensive studies of product market competition and innovation typically restrict attention to manufacturing industries (Aghion *et al.*, 2005, for example).[2]

In what follows, we first discuss previous approaches to aggregate data on technology diffusion and propose to incorporate a time-varying saturation level. We then present our retail industry data and results from country-wise and pooled estimations (section 3). After a discussion of a number of robustness checks, section 4 offers some final remarks.

2 Analytical framework

Patterns of aggregate technology diffusion usually resemble a sigmoid shape. Most empirical studies thus follow Griliches (1957) and employ the logistic function as analytical tool, which captures the sigmoid shape through three easily interpretable parameters:[3]

$$S_t = \frac{\gamma N_t}{1 + \exp(-\beta(t - \tau))}, \tag{1}$$

where S_t denotes the number of technology adopters at time t. Over time, S_t converges to a ceiling or saturation level of adopters, which is a fraction γ of the total population N_t. Timing and speed of this process are determined by parameters τ and β. Provided with data on S_t and N_t and assuming that the three parameters are constant over time, they

[2] Regarding the retail sector, we are only aware of studies based on firm-level data, for example Foster, Haltiwanger and Krizan (2002) and Levin, Levin and Meisel (1987, 1992). Although rich in various aspects, firm-level data typically lack variation in the regulatory environment and hence provide little opportunity to examine policy issues.

[3] With micro-level data discrete choice and hazard rate models are commonly used, for example see Karshenas and Stoneman (1993), Åstebro (2004) and the references therein. For a review see Hall and Khan (2003).

may be estimated; for example by non-linear least squares (NLS) under the assumption of an additive error term.[4]

Properties. For the moment, suppose that β, τ and γ are time-invariant. In that case, the function is symmetric: S_t equals half of its saturation level at the curve's inflection point τ. At date $t = \tau$, the growth rate of the number of adopters is no longer increasing. Hence, τ is a measure for the timing of adoption – it shifts the S-curve forwards or backwards on the timeline. Too see this, consider t^k, the moment in time where a share k of the saturation level is reached:

$$\frac{\gamma N_t}{1+\exp(-\beta(t^k-\tau))} = k\gamma N_t$$
$$\text{or} \quad t^k = \tau - \beta^{-1}\log(k^{-1}-1).$$

At $k = .5$, $t^k = \tau$. Differentiating equation 1 with respect to time shows that coefficient β is a measure for the speed of adoption. It gives the growth rate of S_t, relative to its distance to the saturation level, thereby causing the non-linear shape of the diffusion curve: $\frac{\mathrm{d}S_t}{\mathrm{d}t}\frac{1}{S_t} = \beta\frac{\gamma N_t - S_t}{\gamma N_t}$. The maximum growth rate of S_t is thus $\frac{\beta}{2}$ (attained at time τ).

Most studies, however, use another version of equation 1, where

$$S_t = \frac{\gamma N_t}{1+\exp(-\alpha-\beta t)}.$$

Whereas the advantage of that version is that it lends itself more easily to log-linearization, its disadvantage is that α is erroneously interpreted as a timing indicator. Instead, $\alpha = -\beta\tau$ and hence 'timing' estimates for α resulting from the traditional version are strongly correlated with respective speed estimates for β. It is therefore questionable to interpret this finding as evidence of technological convergence (Gruber and Verboven, 2001).

Reduced-form application. The Griliches approach is to apply the logistic function as a reduced-form tool: much like linear regression, yet accounting for non-linearity in the dependent variable. This approach uses time (t) as a variable to proxy for unobserved factors that underly the diffusion process, for example the decline in the real cost of adoption over time, or slow replacement of working older technology (Griliches, 1980). Another explanation for the fact that technology diffusion is rarely instantaneous is that

[4] Of course, usual econometric suspects such as autocorrelation may have to be dealt with (see section 3.3).

adopters are heterogeneous, for example in their propensity to get informed about the new technology or in their costs or benefits of adoption. Below we argue that, if such population heterogeneity varies over time, this explanation may also involve a time-variant saturation level γ.

A more structural approach is to formulate a theoretical model for the diffusion process in order to generate an estimable diffusion equation. For example, in a seminal paper Mansfield (1961) proposes a model which leads to the logistic function. Yet, since then other models have been proposed that are often hard to distinguish empirically (Geroski, 2000; Stoneman, 2002). Therefore, a reduced-form approach to functional form seems more appropriate especially for studies of macro data, when there is little *a priori* knowledge about the precise data generating process.

Fitness. Still, one may wonder whether the logistic is the right reduced functional form. Based on a large set of data covering numerous countries and technologies, Comin, Hobijn and Rovito (2006) argue that the logistic is not an appropriate functional form to describe the intensive margin in technology diffusion. The intensive margin measures the intensity at which a technology is used across time by its adopters.[5] Most data including ours, however, measure the extensive margin (the number of adopters), for which the logistic seems appropriate. Moreover, Comin, Hobijn and Rovito (2006) note that their findings may be due to the standard assumption of a time-invariant ceiling. Other studies have criticized the symmetry of the logistic function and argued in favor of asymmetric functions such as the Gompertz (Dixon, 1980, for example). In general, and particularly this paper, we prefer Zvi Griliches' position to emphasize data rather than functional form:

> "Adding parameters to the curve itself or fiddling with the functional form is not an attractive alternative, in my opinion. What one gains in fit one loses in interpretability. Instead, I would now respecify the model so that the ceiling is itself a function of economic variables that change over time."
> (Griliches, 1980, p. 1463)

Adding independent variables. For his 1957 paper, Griliches basically estimated the three parameters of equation 1 separately for each of 31 U.S. states, and assessed

[5]The intensive margin corresponding to our data would be the share of retail sales that go through scanner checkouts.

how these estimates correlate with other variables across states. Most subsequent cross-country or cross-industry studies employed a variation of this two-stage approach (see Comin, Hobijn and Rovito, 2006, for a recent example).[6] In this paper, we follow Griliches later suggestion and incorporate independent variables directly into a single-stage cross-country estimation of the logistic function:

$$S_{it} = \frac{(\gamma_i + \mathbf{x_{it}}'\gamma^x)N_{it}}{1+\exp(-\beta_i(t-\tau_i))} \quad , \tag{2}$$

where $\mathbf{x}$ contains a number of independent variables, subscript i indicates countries and t indicates periods.

The coefficients β_i, τ_i and γ_i account for time-invariant country-specific effects as well as for time-invariant cross-country measurement differences. Hence, they serve a similar objective than fixed effects in a standard cross-country panel regression. In other words, we retain the full flexibility of a country-wise estimation of the logistic function but use equation 2 to ask whether the variables contained in $\mathbf{x}$ provide additional information regarding cross-country differences in the long-run diffusion level. Accordingly, γ^x estimates the average marginal effect of variable x on the country-specific saturation level.

Most theoretical studies of technology diffusion, in contrast, focus on explaining diffusion speed. In response, some empirical studies used an approach comparable to equation 2 but relating β to independent variables (Gruber and Verboven, 2001; Koski and Kretschmer, 2004, for example). Such an approach, however, neglects to explain frequently found, considerable cross-country differences in γ. A common response to such findings – and in general to γ-estimates below 1 – is to assert that the population of potential adopters (N) is erroneously specified (Trajtenberg and Yitzhaki, 1989). Specification 2, instead, acknowledges both the inevitable imperfectness of measure N and the time-invariant nature of the long-run diffusion level by allowing γ to be smaller than 1 and subject to time-variant economic factors.

Moreover, from a welfare perspective it is arguably more important to understand drivers of long-run usage rather than diffusion speed of a new technology. In particu-

[6]An alternative approach is a linear cross-country panel regression analysis, in which a potentially non-linear diffusion pattern is partly accounted for by time dummies (Caselli and Coleman, 2001). The respective coefficients are typically assumed to be constant across countries, such that the added independent variables capture cross-country differences in both timing and saturation level of technology diffusion.

lar, population heterogeneity is an important determinant of long-run diffusion patterns, which in consequence are sensitive to changes in heterogeneity over time. Provided $\mathbf{x}$ contains variables that capture variance in heterogeneity, equation 2 is a natural way to estimate the respective effect. Finally, notice that this specification should be more robust in estimations. For example, estimation can encounter convergence problems when the non-linear function to be estimated is too complex. In contrast to the case when $\mathbf{x}$ enters the denominator of the logistic function through β or τ, here it enters the numerator. Equation 2 is thus more linear-like, which facilitates estimation.

3 Application to a retail information technology

Our application is the diffusion of first-generation checkout barcode scanners. The first retail outlet was equipped with a barcode scanner in 1974 in the United States (Nelson, 2001). In Europe, however, diffusion did not take off before the 1980s. Until 1997, the national member organizations of the European Article Numbering Association (EAN) collected data on the number of retail outlets with scanner installations. These data are published for the years 1981 to 1996 in the yearly reports of the EAN.[7] Unfortunately, we were unable to obtain respective U.S. data beyond that presented in table 1; we also lack U.S. data for some of the independent variables. We therefore do not include the U.S. in our econometric analysis. Figure 1 in the appendix plots the dependent variable for the U.K., France, Germany, and Italy.

From today's perspective, the figures in table 1 – but also those for some European countries presented below – appear to exhibit implausibly low adoption rates. After all, our daily shopping experience suggests that barcode scanners are ubiquitous. Yet, notice that we consider not only grocery retailing but the whole retail sector; which includes types of retailers that typically do not use barcode scanners (flower shops, repair shops or bakeries, for example). Furthermore, the EAN data concern only fixed scanners in checkout counters. Many smaller retailers now work with hand-held or mobile barcode scanners, figures on which are not included in our dataset. In our working paper (Beck, Grajek and Wey, 2005), we argue in more detail that – since it is restricted to checkout bar-

[7] The earliest EAN report available (at www.ean-int.org) is the 1983 report, which also gives figures for 1981 and 1982 for most countries (or indicates that there were no scanning stores before 1983 in a particular country).

Table 1: Diffusion of barcode scanning in the U.S., 1974-1984

Year	Scanning stores[a] (number)	Outlets with payroll[b]	Scanning stores (%)	Scanning food stores (%)[c]
1974	6	726940	<.00001	
1976	97	744780	<.00001	
1980	2483	738100	.00003	
1982	5902	784700	.00008	
1984	9278	831300	.00011	
1988				59.7
1989				57.7

Sources:
[a]Euromonitor (1986, based on trade magazine *Chain Store Age*);
[b]Bureau of the Census (1978, and later issues);
[c]Food Marketing Institute (1989,1990).

code scanning in ten industrialized countries – our dataset is sufficient for an econometric study.

Country-wise estimations. Table 2 presents estimates from separate country-wise NLS estimations of equation 1. We also provide the estimations' R^2, yet only to compare fit across countries; high R^2 values are common in such non-linear models and not *per se* suggestive of a good specification (Trajtenberg and Yitzhaki, 1989). The number of outlets, N_t, is counted in hundreds such that $\hat{\gamma}_i$ indicates the estimated saturation level as the percentage of stores with a checkout barcode scanner.

Altogether, country-wise estimation results are in line with the productivity studies cited earlier. Cross-country differences seem to be most pronounced with respect to the saturation level of IT adoption as measured by $\hat{\gamma}_i$. For example, Austria is estimated to have about 24% of outlets with checkout barcode scanning in the long run but Italy only 1%.[8] Differences with respect to timing and speed of diffusion seem less pronounced. Only in two cases do estimates for β_i and τ_i differ significantly from the cross-country average of .42 (which implies a growth rate in the number of barcode scanning stores of 21% around year 1994).

Yet, the estimated absolute values for γ should not be taken too literally: first, measurement of N_{it} may differ across countries; for example, for some countries it may in-

[8]In contrast to the Italian case, we are rather surprised by the low estimated saturation level for Ireland, since Ireland's retail structure is more comparable to that of the U.K. (see table 6 in the appendix). As Ireland has developed strongly throughout the 1990s, we presume that our data cover only the very beginning of a corresponding diffusion process, which complicates estimation (Debecker and Modis, 1994). We return to this point below.

Table 2: Coefficients from country-wise estimations[a]

Country	$\hat{\gamma}_i$	$\hat{\beta}_i$	$\hat{\tau}_i$	Observations	R^2
Austria	24.2*	.50*	1994.0	14	.999
Belgium	16.0*	.39	1994.1	12	.999
Denmark	10.7	.42	1992.1	15	.992
France	10.7	.41	1994.4	13	.996
Germany	5.2*	.41	1992.7*	15	.999
Ireland	1.3*	.48	1992.7*	16	.998
Italy	1.1*	.45	1992.0	15	.986
Netherlands	7.6	.31*	1994.8	14	.997
Spain	3.9*	.39	1995.1	16	.978
United Kingdom	15.4	.41	1995.8	16	.995
Cross-country average	9.6	.42	1993.8		

[a]Parameter estimates from country-wise NLS estimation of equation 1. Starred coefficients differ significantly from cross-country average (95% confidence level, F-test based on asymptotic standard errors).

clude mobile outlets (street traders), for others not.[9] Second, and more importantly, differences in adoption patterns may arise from differences in population heterogeneity, that is, underlying retail market structures. The Italian retail market, for example, is highly segmented, with many small but specialized retailers selling goods that in other countries are sold jointly by larger retailers. Like most other countries – but eventually to a different degree – Italy experienced considerable changes in retail market structure in the last decades. Such changes in heterogeneity not necessarily affect the absolute value of N_{it}, but rather its composition and consequently long-run levels of IT diffusion estimated by $\hat{\gamma}$.

We conclude that our application constitutes a well-suited case for an econometric approach to relate differences in γ to independent variables that capture retail heterogeneity as well as other economic factors. With the country-specific coefficients γ_i in equation 2, the pooled estimation accounts for time-invariant cross-country as well as potential measurement differences in estimating the marginal effect of independent variables on long-run diffusion.

[9]Measurement differences do not appear to be substantial, however: we obtain similar cross-country differences if we relate the number of barcode scanning stores to population instead of the total number of outlets (see section 3.3).

Table 3: Description of independent variables

Label	Description	Source	Cross-country mean 1981 / 1996
OUT	No. of retail outlets (per mn. inhabitants)	Euromonitor, World Bank	9361.8 / 7952.4
HYP	No. of hypermarkets (per mn. inhabitants)	Euromonitor, World Bank	6.8 / 13.3
EPL	OECD indicator of strictness of employment protection legislation	OECD	2.5 / 2.2
WAGE	Retail hourly real wage (index 1995=100)	GGDC, World Bank	74.2 / 101.1
GDP	Per capita real GDP (index 1995=100)	World Bank	74.8 / 102.1
VOL	Retail sales volume (index 1995=100)	OECD, Euromonitor	85.7 / 101.3

3.1 Independent variables

A complicating issue for empirical studies of the retail sector is that publicly available information is scarce, even on the country-year level. Although we compile data from various sources, a number of limitations make us restrict attention to 10 European countries. Table 3 describes the corresponding set of independent variables (table 6 in the appendix gives more detailed summary statistics).

Data on the total number of retail outlets (*OUT*) and the number of hypermarkets (*HYP*) are from various issues of "Retail trade international", a publication by market research firm Euromonitor. The most recent issue is Euromonitor (2002). Source for *GDP* and population figures is the World Bank (2003). As a measure for the severity of labor market restrictions, we use version 1 of the revised OECD indicator of the strictness of employment protection legislation (OECD, 2004). The indicator of retail sales volume (*VOL*) is also from the OECD.[10] The retail *WAGE* index is constructed using data from the 60-Industry Database of the Groningen Growth and Development Centre (GGDC). Some missing values had to be replaced with univariate procedures (the appendix in Beck, Grajek and Wey, 2005, gives a detailed list of all data manipulations). We were unable to obtain information on a number of factors that may also be important in our analysis, such as prices for scanning equipment, opening hours, the importance of multinational

[10]For Italy and Spain, this indicator does not cover the whole sample period. For these two countries, we therefore constructed a comparable indicator based on Euromonitor and GGDC data (see Beck, Grajek and Wey, 2005, for more details).

firms, or average store size. As long as these omitted factors are relatively time-invariant or equal for the countries in our sample, they are accounted for by the country-specific estimates for γ_i.

Predicted effects. When deciding about the adoption of a new technology, a firm typically compares costs and benefits of adoption at a given point in time (Hall and Khan, 2003). In our case, the installation of a barcode scanner represents a major capital investment that basically enables a retailer to check out more retail items in less labor time.[11] Following the discussion by Levin, Levin and Meisel (1987, 1992), our independent variables capture a number of factors that can make barcode scanning more or less valuable in different countries.[12] First, the expected financial returns to such a capital investment depend on future market conditions. Since return-on-investment is quicker in growing markets, retailers there will adopt more intensely than retailers in stagnating or contracting markets. In addition, barcode scanning may introduce or increase economies of scale in retailing. In both cases, we expect adoption intensity to increase with market volume (*VOL*). Second, barcode scanning is likely to reduce customer waiting time at the checkout. Customers in high-income countries have a higher opportunity cost of waiting. Using per-capita *GDP* as income measure, we expect diffusion of barcode scanning to increase with *GDP*. In this interpretation, barcode scanning is a product-enhancing innovation: it increases the quality of retailing for the customer.

Another more classical interpretation regards barcode scanning as a process-enhancing innovation. Most prominently, barcode scanning may be a labor-saving technology that reduces total labor demand. In addition to this classic capital-labor substitution effect, barcode scanning may allow retailers to substitute unskilled for costly skilled labor. Clerks at scanner checkouts need neither know prices nor be able to type quickly. In both cases of substitution, we thus expect countries with rising retail wages (*WAGE*) to invest more in a labor-saving technology such as barcode scanning. In contrast, strict employment protection legislation (*EPL*) may prohibit retailers from substituting barcode scanners for labor as extensively as the technology might allow (IMF, 2001).

[11] Clearly, barcode scanning also facilitates other potentially productivity-enhancing practices, e.g. sophisticated logistics systems ('efficient consumer response', 'category management'); but these systems did not develop before the mid-1990s and still seem to represent "untapped potential" (Haberman, 2001).

[12] Levin, Levin and Meisel (1987, 1992) study the adoption of barcode scanning in U.S. retailing. They analyse firm-specific data relating to the early years of the technology (1974-1985).

Retail competition. Our particular focus, however, is on the role of competition. In the industrial organization literature, retail markets have long been regarded as more or less perfectly competitive. This perception has led scholars to abstract from the retail level and concentrate on the manufacturers' side. Yet, fragmented retail structures are most often the direct result of entry restrictions. In general, these restrictions tend to favor small retailing in downtown areas against large scale retail formats as exemplified by *Wal-Mart*. Most prominently, planning and construction restrictions have been used in all European countries to ban large retailing formation; e.g., by not granting construction permissions or by limiting store size (Bertrand and Kramarz, 2002). Planning and construction restrictions have been eased first in the U.K. by the Thatcher government and later in other European countries as well.

With these developments, hypermarkets have become an integral element of most retail markets. According to the standard definition, hypermarkets have a minimum size of 2,500 square meters, and sell both food and non-food items.[13] Hypermarkets often locate in peripheral areas which are easily accessible by car; in the U.S. similar stores are often called "superstores". In most European countries, the hypermarket retail format emerged in the 1970s and 1980s.[14] In this paper, we use the number of hypermarkets per capita (*HYP*) as an inverse indicator of entry restrictions: an increasing number of hypermarkets is a result of less restrictive entry regulations, and hence a proxy for increasing competitive intensity due to regulatory change. Moreover, hypermarkets may reflect competitive intensity on other grounds. They can be regarded as low-cost competitors who exploit the cost benefits of out-of-town locations, sophisticated logistics, and economies of scale (Basker, 2005). One may also view retail competition as competition of retail channels or formats (Smith and Hay, 2005). In that sense, the emergence and growth of a new format, like the hypermarket, intensifies retail competition as such.

In our working paper (Beck, Grajek and Wey, 2005), we present precursory evidence supporting this view. Namely, the appearance of hypermarkets seems to lead to increased market exit rates of other retail outlets. With regard to IT diffusion, such hypermarket

[13]Two countries in our sample – Germany and Denmark – apply a slightly broader hypermarket definition which includes large supermarkets with a floor space between 1,500 and 2,500 square meters. In our pooled estimation below, we allow for a different hypermarket effect for these two countries.

[14]*Carrefour*, one of the world's largest retailers, claims to have invented the concept. It opened its first hypermarket in 1963 near Paris, "with a floor space of 2,500 square meters, 12 checkouts and 400 parking spaces" (see www.carrefour.com/english/groupecarrefour/annees60.jsp).

competition can have two – potentially independent – effects.[15] On the one hand, since hypermarket entry seems to induce exit of other retailers, this may entail a *selection effect*: if the exiting retailers predominantly belong to the group of (long-run) IT adopters, hypermarket competition reduces the share of adopters in the remaining retailers. On the other hand, it can have an *encouragement effect* on the remaining retailers, for example when hypermarket entry leads former non-adopters to become adopters in the long run. With industry-level data on IT diffusion, we can only identify the joint impact of these two effects, which can be positive or negative.

For an illustration, consider a simple numerical example. Imagine a country with 100 retailers, 50 of which are potential adopters of barcode scanning. There are no hypermarkets yet. While barcode scanning diffuses, one of the retailers decides to restructure and become a hypermarket, which drives 10 other retailers out of business. Depending on whether these 10 quitting retailers were potential adopters or not, the *selection effect* of the hypermarket on long-run diffusion can be negative or positive. In case all quitters were non-adopters, the long-run diffusion level increases from .5 to .56 (50 out of 90). In case they had been (potential) adopters, it reduces to .44 (40 out of 90). Moreover, an *encouragement effect* of increased competition could be that some of the previous non-adopters become potential adopters of barcode scanning, which raises its long-run diffusion level.

Bivariate correlations. For a first idea on how the discussed factors might relate to cross-country differences in barcode scanning, we first follow the standard two-stage approach and assess how the countries' separately estimated saturation levels correlate with country-specific trends of the proposed variables. Table 4 lists the correlation coefficients.

All bivariate correlation coefficients are in line with the above theoretical expectations, except for EPL. Estimated saturation levels are higher in countries with larger growth of GDP, retail sales volume, retail wages and employment protection, and lower in countries with larger hypermarket growth.[16] A negative hypermarket effect is surprisingly clear in the data: between 1981 and 1996, 5 out of 10 countries have an average yearly growth in the per-capita number of hypermarkets below 3% – as proxied by

[15] In section 3.3, we argue that reverse causality or endogeneity are not affecting the observed relationship between hypermarkets and IT diffusion.

[16] When we include Ireland in calculating these correlation coefficients, only the coefficient for GDP changes qualitatively, resulting from Ireland's combination of strong GDP growth with a low γ-estimate.

Table 4: Bivariate correlations between $\hat{\gamma}_i$ and independent variables[a]

Correlation between					
		country-specific trend coefficient for			
	log(HYP)	log(EPL)	log($WAGE$)	log(GDP)	log(VOL)
and $\hat{\gamma}_i$	-.526	.512	.501	.015	.581

[a]Based on nine observations (one per country, excluding Ireland): $\hat{\gamma}_i$ and trend coefficient from country-wise regression of log(x_{it}) on t. Trend coefficients are significant with 95% confidence for all countries and variables except for two countries with variable VOL.

a trend coefficient in a regression of log(HYP). The average estimated saturation level is 12.7% for these countries (Austria, Belgium, Denmark, Germany, Netherlands), but only 6.5% for the other five countries that had stronger hypermarket growth. Yet, these bivariate correlations neither account for time-invariant country-specific effects in γ – eventually affected by measurement differences – nor for year-to-year and multivariate correlations.

3.2 Pooled estimation

In order to account for these effects, we pool countries and estimate a joint logistic function (equation 2).[17] Estimations based on the full sample, however, exhibited convergence problems and led to large and unstable estimates for Ireland's country-specific coefficients γ_i, β_i and τ_i (unreported). We actually find this result reaffirming in two respects. First, it seems to confirm the suspicion that the data for Ireland do not cover a sufficiently large portion of its diffusion of barcode scanning. Second, it suggests that the independent variables do contain additional information, since Ireland-specific estimates based on equation 1 (thus excluding $\mathbf{x}$) spuriously appeared stable. In what follows, we therefore present results excluding Ireland. The independent variables' coefficients are similar to estimates including Ireland, but convergence is smoother and all country-specific estimates are now stable.

The first column of Table 5 presents the results for our baseline specification (I), where D is a dummy variable equal to one for Germany and Denmark to account for the different, somewhat broader hypermarket definition employed for these countries. In

[17]We use the estimates from the country-wise regressions as initial values for country-specific effects. For the independent variables' coefficients, we set initial values equal to 0.

Table 5: Results for pooled estimations[a]

Dependent variable: Number of barcode scanning stores			
Specification	(I)	(II)	(III)
HYP	-1.852*	-1.744*	-1.756*
	(.426)	(.408)	(.425)
*D*HYP*	6.668*	1.837	
	(2.505)	(2.154)	
EPL	-1.333	-1.154	
	(2.287)	(2.108)	
WAGE	.119*	.123*	.116*
	(.032)	(.031)	(.028)
GDP	.394*	.385*	.408*
	(.069)	(.069)	(.061)
VOL	.087§	.072	.062
	(.048)	(.047)	(.047)
Country excluded:	Ireland	Ireland Germany	Ireland Germany
Time span (max.)	1981-1996	1981-1996	1981-1996
Observations	130	115	115
Adj. R^2	.994	.994	.994
Root MSE	494.6	504.2	499.4

[a]Estimates for γ_i, β_i and τ_i omitted. Asymptotic standard errors in parentheses. Significance levels: *95%, §90%

addition, we present results for two alternative specifications: in specifications II and III we exclude Germany; in specification III, we also exclude the variables *EPL* and *D*HYP*.

The effects for variables *WAGE*, *GDP* and *VOL* vary little across specifications and are for the most part significantly estimated. A 10-point increase in the retail wage index is associated with an estimated increase in the saturation percentage of barcode scanning stores by about 1.2 points on average. A 10-point increase in real GDP per capita and retail sales volume is associated with an estimated increase in the saturation level by about 4 and 1 points, respectively. All three results confirm initial expectations: First, investment in labor-saving retail IT can be interpreted as a reaction to changes in labor costs. Second, income, scale and returns-to-investment effects are important. Although the income effect measured by *GDP* seems more important than the scale effect measured by *VOL*, the effects are hard to distinguish empirically since the two variables are by definition correlated.

Consistent with the bivariate correlation found before, an increase in the number of hypermarkets by one per million inhabitants is estimated associated with an estimated decrease in the saturation percentage of barcode scanning stores by almost 2 points. In aggregate terms, hypermarket competition therefore seems to reduce long-run IT usage in the retail sector. The question whether this hypermarket effect works by discouraging existing retailers from adoption (*encouragement effect*) or rather by driving potential adopters out of the market (*selection effect*) is one we cannot address with the data at hand. We suspect that both effects are at play. The impact of the selection effect may be more important, however, since hypermarkets mainly compete with (and induce exit of) supermarkets – the main group of potential adopters – and less with other, smaller retailers.

Yet, the negative result appears to hold only for the standard hypermarket definition. In our baseline specification I, the estimate for the broader Danish/German definition is positive (-1.9+6.7). German figures, however, may be affected by exceptional efforts due to reunification: from 1989 onwards, East Germany experienced a catch-up in the number of retail outlets, many of which equipped with barcode scanners from the start. We therefore re-estimated the model excluding Germany and find that Germany indeed seems to be a special case. The estimated difference between the broad hypermarket definition and the standard one, now a Denmark-specific effect, is much lower and not significantly different from zero. Accordingly, an estimation which ignores different hypermarket definitions by excluding the interaction term *D*HYP* (specification III) leads to an essentially unchanged hypermarket effect, as long as Germany remains excluded. In specification III we also exclude the *EPL* indicator, whose effect has the expected negative sign but is insignificant in all estimations. Other estimates remain largely the same.

3.3 Robustness

Autocorrelation can be an issue in estimating growth curves. Following a relatively straightforward testing procedure (Franses, 2002), in country-wise estimations we reject the null hypothesis of no autocorrelation against the alternative of AR(1) errors only for two countries (results available upon request). Yet, re-estimating a logistic function with an AR(1) error term for these countries leads to autocorrelation coefficients which are not significantly different from zero. We therefore retain the assumption of an AR(0) error term throughout the paper.

Endogeneity and reverse causality. A potential source of endogeneity bias is the presumption that every new hypermarket built from the mid-1980s increases the number of scanning outlets by one. Although not necessarily, since hypermarkets operated long before the introduction of barcode scanning and hence the technology may not be as crucial for them as it might appear from today's perspective. In any case, the negative estimates in table 5 already suggest that this endogeneity bias cannot be very influential. By deducting the number of hypermarkets from both the number of barcode scanning stores and the number of outlets, it is nevertheless possible to focus on the effect of hypermarket competition on the adoption of barcode scanning by all other retailers. The corresponding (unreported) results are virtually identical to those in table 5.

Moreover, one may suspect that there are effects driving reverse causality, namely that barcode scanning leads to an increase in average store size and eventually to more "superstores" or hypermarkets (Holmes, 2001). Yet, the facts that (*i*) hypermarkets existed long before barcode scanning was introduced and that (*ii*) Holmes (2001) predicts a positive correlation while we find a negative one suggest that reverse causality is not a severe issue in our data set.

Data concerns. Another potential source of error are the implicit assumptions in our method to construct time series for the total number of retail outlets (see the data appendix in Beck, Grajek and Wey, 2005). We therefore estimated specifications I to III with a country's population (in millions) replacing the number of outlets in equation 2. For specifications II and III, all variables yield estimates with the same qualitative effects on the long-run number of barcode scanning stores per capita; except EPL, whose coefficients change sign but are again insignificant. Only for specification I, which includes the special case of Germany, some results differ (reported in Beck, Grajek and Wey, 2005). We infer that our results are not crucially affected by the data manipulations that were necessary to obtain a workable time series for the number of retail outlets.

Finally, our conclusions regarding the effect of *EPL* may be premature. Given missing values and other measurement problems associated with the OECD *EPL* index (Blanchard and Wolfers, 2000), there are reasons to doubt the validity of the indicator used. In order to cross-check results, we replaced the *EPL* indicator with country time series on the cumulative number of EPL reforms from the *Social Reforms Database* of the Fondazione Rodolfo DeBenedetti. When replaced for our initial *EPL* indicator in specification II, these variables also yield insignificant results (available upon request).

4 Concluding remarks

The first generation of checkout barcode scanners, a critical information technology in retailing, has diffused to different levels of long-run usage across European countries. Although similar observations hold for other technologies (Comin, Hobijn and Rovito, 2006), cross-country studies typically focus on explaining differences in diffusion speed or timing. In this paper, we propose a novel, single-stage econometric approach to relate long-run usage to independent factors within the standard reduced-form model of technology diffusion. This approach can be fruitful in other applications, particularly when heterogeneity in the population of technology adopters varies across observations.

Results for our application are consistent with earlier studies. We find that, as expected, retail IT diffusion is more intense in countries with large and growing retail markets and economies. It is therefore not surprising that the United States, which experienced strong overall economic growth driven by a surge in consumer spending, are ahead of most European countries when it comes to IT diffusion in the retail sector and the resulting productivity gains throughout the 1990s.[18] In line with standard theory, we also find that raising labor costs induce retailers to substitute barcode scanners for labor. In contrast, we do not find employment protection legislation to significantly impact retail IT diffusion. With respect to an upcoming “retail revolution” that relies on RFID technology, our results lead us to expect stronger RFID diffusion in countries that exhibit scale effects and upward wage pressure.

With regard to retail competition, our results suggest that the rise of hypermarkets reduces long-run retail IT diffusion. This effect, which is robust in a variety of specifications, has two potential explanations. First, hypermarket competition causes exit of potential IT adopters, namely smaller-sized supermarkets. Second, hypermarkets – which are most likely to adopt barcode scanning early – discourage subsequent adoptions of

[18]Comparable OECD data for the retail volume indicator *VOL* indicate that, between 1990 and 2000, U.S. retail volume increased by about 67%, whereas it increased by about 30% in the U.K. and by about 7% in France. In Germany, retail volume decreased by about 1% between 1990 and 2000.

rival retail formats. Overall, these results suggest that liberalization of retail market entry and the associated emergence of hypermarkets deepens retail segmentation such that hypermarkets on the one hand and small down town retailing (including shopping mall retailing) prevail. In contrast, intermediate retail formats – in particular medium-sized supermarkets – are likely to suffer from market liberalization.

Yet, as we do not observe the intensive margin of retail IT diffusion, the productivity implications of our findings are not evident. Depending on how much retail volume goes through barcode scanning retailers, IT productivity may increase in relative terms with the emergence of hypermarkets, even though aggregate IT usage decreases at the extensive margin. Also, our data are not directly comparable to official measures of retail IT investments, since they count the number of barcode scanning stores, not the number of scanner installations. In our data, a smaller supermarket with, say, one scanner checkout has the same weight as a larger one with multiple scanner checkouts. Further research may include more countries and explanatory variables, for example measures of foreign direct investments in order to assess the role of large multinational retail firms in IT diffusion. Reviewing our efforts to put together the present dataset, we however fear such a task is more demanding than it seems at first sight.

A Appendix

Figure 1: Number of barcode scanning stores (in %, by country)

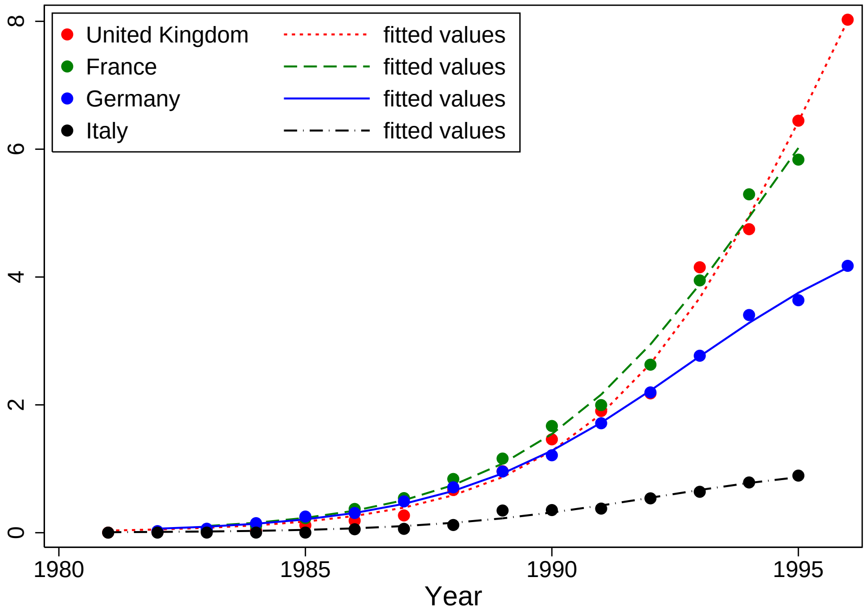

Table 6: Detailed summary of variables[a]

Variable	*OUT*	*HYP*	*EPL*	*WAGE*	*GDP*	*VOL*
Country						
Austria	4762.5	29.9	2.2	87.9	94.0	95.2
	356.2	15.6	.2	45.3	37.6	31.8
Belgium	4653.9	7.5	2.8	87.8	93.9	94.8
	677.1	1.7	.9	37.2	36.6	42.9
Germany	4652.7	22.7	2.9	91.9	93.9	94.2
	1937.8	11.6	.9	37.0	31.5	27.2
Denmark	6841.3	16.0	1.8	90.3	94.7	96.5
	896.8	8.0	.7	49.0	34.7	23.2
Spain	20730.4	4.0	3.4	93.7	92.8	100.0
	8573.0	6.6	.8	43.1	46.0	27.3
France	7159.5	15.3	2.8	93.6	95.5	98.4
	2380.6	11.7	.3	30.1	33.3	17.1
Ireland	9177.1	5.3	.9	92.7	91.1	100.0
	707.9	13.1	.1	68.3	99.0	74.5
Italy	16139.8	3.9	3.6	94.7	93.4	99.0
	6221.5	9.1	1.9	17.0	33.7	25.0
Nether-	5469.4	2.4	2.5	95.8	94.3	101.5
lands	652.0	1.5	.6	24.0	40.2	25.0
United	6740.1	3.4	.5	91.0	93.1	93.6
Kingdom	2564.8	4.6	.2	40.8	44.8	61.5

[a]Country-specific means in the first line, in the second line the difference between the maximum and the minimum value observed in the respective series (range).
See table 3 for a description of the variables.

References

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith and Peter Howitt (2005), Competition and innovation: An inverted-U relationship, *The Quarterly Journal of Economics*, 120(2): 701–728.

Ark, Bart van, Robert H. McGuckin and Matthew Spiegelman (2005), The retail revolution – can Europe match U.S. productivity performance?, Report 1358, The Conference Board.

Åstebro, Thomas (2004), Sunk costs and the depth and probability of technology adoption, *Journal of Industrial Economics*, 52(3): 381–399.

Basker, Emek (2005), Selling a cheaper mousetrap: Wal-mart's effect on retail prices, *Journal of Urban Economics*, 58(2): 203–229.

Beck, Jonathan, Michal Grajek and Christian Wey (2005), Hypermarket competition and the diffusion of retail checkout barcode scanning, CEPR Discussion Paper 5386.

Bertrand, Marianne and Francis Kramarz (2002), Does entry regulation hinder job creation? Evidence from the French retail industry, *The Quarterly Journal of Economics*, 117(4): 1369–1413.

Blanchard, Olivier and Justin Wolfers (2000), The role of shocks and institutions in the rise of European unemployment: The aggregate evidence, *The Economic Journal*, 110(462): C1–33.

Bureau of the Census (1978), *Statistical abstract of the United States: 1978*, United States Bureau of the Census.

Caselli, Francesco and Wilbur J. Coleman (2001), Cross-country technology diffusion: The case of computers, *The American Economic Review*, 91(2): 328–335.

Comin, Diego, Bart Hobijn and Emilie Rovito (2006), Five facts you need to know about technology diffusion, NBER Working Paper.

Debecker, Alain and Theodore Modis (1994), Determination of the uncertainties in S-curve logistic fits, *Technological Forecasting & Social Change*, 46: 153–173.

Dixon, Robert (1980), Hybrid corn revisited, *Econometrica*, 48(6): 1451–1461.

Euromonitor (1986), *Retail Trade International: 1986 Edition*, Euromonitor International Plc., London.

— (2002), *Retail Trade International: 2002 Edition*, Euromonitor International Plc., London.

Food Marketing Institute (1989), *The food marketing industry speaks, 1989. Detailed tabulations*, Food Marketing Institute.

— (1990), *The food marketing industry speaks, 1990. Detailed tabulations*, Food Marketing Institute.

Foster, Lucia, John Haltiwanger and C. J. Krizan (2002), The link between aggregate and micro productivity growth: Evidence from retail trade, NBER Working Paper Series, No. 9120.

Franses, Philip Hans (2002), Testing for residual autocorrelation in growth curve models, *Technological Forecasting & Social Change*, 69: 195–204.

Geroski, Paul A. (2000), Models of technology diffusion, *Research Policy*, 29: 603–625.

Griliches, Zvi (1957), Hybrid corn: An exploration in the economics of technological change, *Econometrica*, 25(4): 501–522.

— (1980), Hybrid corn revisited: A reply, *Econometrica*, 48(6): 1463–1465.

Gruber, Harald and Frank Verboven (2001), The diffusion of mobile telecommunications in the European Union, *European Economic Review*, 45: 577–588.

Haberman, Alan L. (ed.) (2001), *Twenty-five years behind bars*, Harvard University Press.

Hall, Bronwyn H. and Beethika Khan (2003), Adoption of new technology, in: Jones, Derek C. (ed.), *New Economy Handbook*, Academic Press.

Holmes, Thomas J. (2001), Bar codes lead to frequent deliveries and superstores, *RAND Journal of Economics*, 32(4): 708–25.

IMF (2001), *The Information Technology Revolution*, chap. III, 105–144, International Monetary Fund.

Karshenas, Massoud and Paul L. Stoneman (1993), Rank, stock, order, and epidemic effects in the diffusion of new process technologies: An empirical model, *RAND Journal of Economics*, 24(4): 503–528.

Koski, Heli and Tobias Kretschmer (2004), Entry, standards and competition: Firm strategies and the diffusion of mobile telephony, *Review of Industrial Organization*, 26(1): 89–113.

Levin, Sharon G., Stanford L. Levin and John B. Meisel (1987), A dynamic analysis of the adoption of a new technology: the case of optical scanners, *The Review of Economics and Statistics*, 69: 12–17.

— (1992), Market structure, uncertainty, and intrafirm diffusion: The case of optical scanners in grocery stores, *The Review of Economics and Statistics*, 74(2): 345–50.

Mansfield, Edwin (1961), Technical change and the rate of imitation, *Econometrica*, 29(4): 741–766.

Nelson, John E. (2001), Scanning's silver celebration, in: Haberman (2001), 25–33.

OECD (2004), *Employment outlook*, Organisation for Economic Co-operation and Development.

Scarpetta, Stefano, Philip Hemmings, Thierry Tressel and Jaejoon Woo (2002), The role of policy and institutions for productivity and firm dynamics: evidence from micro and industry data, Tech. rep., OECD Economics Department.

Smith, Howard and Donald Hay (2005), Streets, malls, and supermarkets, *Journal of Economics & Management Strategy*, 14(1): 29–59.

Stiroh, Kevin J. (2002), Information technology and the U.S. productivity revival: What do the industry data say?, *American Economic Review*, 95(5): 1559–1576.

Stoneman, Paul (2002), *The economics of technological diffusion*, Blackwell, Oxford.

Trajtenberg, Manuel and Shlomo Yitzhaki (1989), The diffusion of innovations: A methodological reappraisal, *Journal of Business & Economic Statistics*, 7(1): 35–47.

World Bank (2003), *World development indicators CD-ROM*, World Bank.

ESMT Working Papers (2007)

	ESMT No.	Competence Center
Estimating Network Effects and Compatibility in Mobile Telecommunication Michał Grajek, ESMT	**07-001**	**European Competitiveness**
Estimating Level Effects in Diffusion of a New Technology: Barcode Scanning at the Checkout Counter Jonathan Beck, Humboldt Universität zu Berlin Michał Grajek, ESMT Christian Wey, Technische Universität, Berlin	**07-002**	**European Competitiveness**

ESMT
European School of Management
and Technology GmbH

ESMT Campus
Schlossplatz 1
10178 Berlin
Phone: +49 (0) 30 212 31-1279

www.esmt.org