Paper for National Bureau of Economic Research, Conference on Research in Income and Wealth "Producer Dynamics: New Evidence from Micro Data, April 8 and 9, 2005 Bethesda, MD

Entry, Exit and Labour Productivity in U.K. Retailing: Evidence from Micro Data*

Jonathan Haskel
Queen Mary, University of London; AIM, CeRiBA, CEPR and IZA

Raffaella Sadun
Centre for Economic Performance, London School of Economics, AIM and CeRiBA

March 2005

Abstract

We investigate UK productivity growth in retailing using business level data 1998-2002. We set out the data available, document productivity levels and productivity growth. We find evidence of considerable productivity dispersion. We also find that adjusting for part-timers affects level comparisons but not in the time series. We also find that entry and exit was a substantial contributor to productivity growth over the period.

*Contact: Jonathan Haskel, Queen Mary, University of London, Economics Dept, London E1 4NS, j.e. haskel@qmul.ac.uk.
Financial support for this research comes from the ESRC/EPSRC Advanced Institute of Management Research, grant number RES-331-25-0030 and is carried out at CeRiBA, the Centre for Research into Business Activity, at the Business Data Linking Branch at the ONS; we are grateful to all institutions concerned for their support. This work contains statistical data from ONS which is crown copyright and reproduced with the permission of the controller HMSO and Queen's Printer for Scotland. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. We thank Ralf Martin for helpful discussions and Felix Ritchie (ONS) for help on the data. Errors are our own.

1 Introduction

Retailing productivity has become an area of acute UK policy interest in recent years, for at least two reasons. First, much productivity analysis is of manufacturing and it is therefore of interest to study productivity in the service sector. Second, in Basu et al's (2003) study of the US TFP acceleration in the late 1990s versus the UK deceleration, they find that retail trade and hotels and catering "account" for about three-quarters of the U.S. acceleration (Domar weighted industry TFP growth) and one-third of the U.K. deceleration. Retailing therefore looks to be a potentially crucial sector.

Our purpose in studying retailing at the micro level is threefold. First, we are not aware of any UK micro level studies of the entire sector before.[1] Second, Foster, Haltiwanger, and Krizan (2002) find that productivity growth in US retailing is due to the entry and exit of new stores, rather than productivity growth in incumbent stores. This has generated a good deal of policy interest since it has been conjectured that restrictive UK planning and/or land use regulations might hold UK productivity back in this crucial sector. One test of this is to see how much productivity growth in UK retailing arises from entry and exit and this can only be done using the micro data. Third, and related, there is some evidence of increasing returns to scale in retailing (Competition Commission, 2002) and thus possible restrictions on entry and exit might be crucial.

The paper proceeds as follows. In section 2 we document the data sources, then describe, in section 3, entry and exit. Section 4 looks at productivity levels and growth. Section 5 concludes.

2 The retailing data

2.1 Time period and industries.

The data in this paper comes from the Annual Respondents Database (ARD). This is a comprehensive business database that is based on the Annual Business Inquiry (ABI) performed by the Office for National Statistics (ONS). Regarding time period, the data available to us is annual from 1997. As we shall see however the 1997 data is not accurate and therefore in practice our analysis starts in 1998. At time of writing the 2002 data was the final period available.[2]

[1] With the exception of Haskel and Kwanja (2001), an early version of this paper. The main difference between this paper and the previous one is that this one uses and extra year of data, and computes numbers using a different employment measure. The latter turns out to make a substantial difference since the earlier employment measure was available only for a subset of firms causing many firms to be dropped. This affects the productivity decompositions, see below.

[2] We particularly thank Felix Ritchie for helping in the timely provision of the 2002 data.

As for industries, the ARD database covers almost all firms with Standard Industrial Classification (SIC) codes from 2010 to 93050. The retailing sector is covered by SIC92 codes from 52111 to 52740 i.e. all codes beginning with 52. Retailing is then split into 7 broad categories, as listed in Table 1.

2.2 Units of analysis

A crucial issue in what follows will be whether the analysis is by store, chain of stores, or chain of chain of stores. This section sets out in some detail what data are available to us.[3] To summarise,

(a) Employment, entry and exit data are available by store, but

(b) Productivity data are available by chain of stores.

2.2.1 Business structure: enterprises, enterprise groups and local units

The fundamental business data set in the UK is the Interdepartmental Business Register (the IDBR). This business register is compiled using a combination of tax records on VAT and PAYE, information lodged at Companies House, Dun and Bradstreet data and data from other surveys. The IDBR has been operating since 1994 (before the IDBR register information was rather uncoordinated across different government departments). The IDBR tries to capture the structure of ownership and control of firms and plants or business sites that make up the UK economy using 3 aggregation categories: local units, enterprises and enterprise groups. Their meaning is best illustrated by means of an example set out in Figure 1.

Consider the left hand panel. Suppose that *Brown* is a single business, operating in a single location, producing goods for a single industry. Now consider the right side of the panel. *Smith and Jones Holdings* are a holding company, registered in London. In turn, they own two businesses, *Smith* and *Jones*, who are involved in separate industrial activities. Smith has four shops (or more generally plants/business sites i.e. a particular geographic location where trade occurs), *Smith North*, *Smith South*, *Smith East* and *Smith West*. *Jones* has a shop, *Jones North* and an R&D lab, *Jones R&D*.

Brown, being responsible for a single business activity, is an "enterprise". *Smith and Jones Holdings*, owing businesses with distinct business activities, is called an "enterprise group"[4]. *Smith* and *Jones* are two enterprises.

What of the shops? All business sites, a business entity at a single mailing address, are called "local units". Consequently if *Jones R&D* is located at a different site than *Jones North* the enterprise *Jones* would consist of two local units. If *Jones R&D* was located at the same site

[3] It follows closely Criscuolo, Haskel and Martin (2003).

[4] A holding company responsible for a number of enterprise groups is called an "apex enterprise".

as *Jones North* the two would form one local unit for the IDBR[5]. (The diagram also refers to reporting units, this will be explained below).

2.2.2 Maintaining information on business structure: enterprise groups, enterprises and local units

The Annual Register Inquiry (ARI) is designed to maintain the business structure information on the IDBR (Jones, 2000, p.51). It began operation in July 1999 and is sent to large enterprises (over 100 employees) every year, to enterprises with 20-99 employees every four years and to smaller enterprises on an *ad hoc* basis. The ARI currently covers around 68,000 enterprises, consisting of about 400,000 local units. It asks each enterprise for employment, industry activity and the structure of the enterprise. For the *Brown* enterprise in our example this is straightforward. A multi-site enterprise such as *Smith* receives a form and is asked to report on its overall activity and employment. It will also be sent 4 extra forms to report the same for each local unit. If *Smith* has closed a local unit it must report this on the form. If a local unit has opened it has to fill out extra forms, which are obtained from ONS by an automated procedure. Returns from the ARI update the IDBR in the summer of each year.

2.2.3 Maintaining information on employment, turnover and other data

As well as the structure of business information, the IDBR holds other data, such as address and SIC code. However, since the IDBR is based mostly on tax data (plus old records from previous inquiries), it also sometimes contains other data. Output information on the IDBR comes from VAT records if the original source of business information was VAT data. Employment information comes from PAYE data if that is the source of the original inclusion. Thus as long as the single-local unit enterprise *Brown* is large enough to pay VAT ((the threshold was £52,000 in 2000/01) it would have turnover information at the enterprise and local unit level. On the other hand, if *Brown* does not operate a PAYE scheme, it will have no employment information. However, employment data is required to construct sampling frames and hence is interpolated from turnover data. For the multi-local unit enterprise *Smith*, no turnover information will be available for *Smith's* local units, since most multi-local unit enterprises do not pay VAT at the local unit level. If the PAYE scheme is operated at the local unit level, it would have independent employment data.

There are two other ways in which employment and output data are gathered. The first is if the business is included ARI (although note that the and the second if it is included in the Annual Business Inquiry (ABI), see below.

[5] The two could nevertheless be separate local units depending on the survey. If for example an R&D survey which collects data just for the R&D part of the business was undertaken, this would identify them as distinct. Thus some care has to be taken in matching business using different surveys.

2.3 The ABI and the ARD.

Whilst the IDBR holds much useful information, more data is required on outputs and other inputs, in order to calculate GDP. Thus the ONS conducts a business survey, based on the IDBR. This is the Annual Business Inquiry (ABI) and the ARD consists of the panel micro-level information obtained from successive cross-sections of the ABI. The ABI covers production, construction and some service sectors, but not public services, defence and agriculture.[6]

The questions asked on the ABI for retailing vary somewhat. They are required to provide details on turnover (total and broken down in retail and non-retail components, and by commodity sold), expenditures (employment costs, total materials and taxes), items defined as work in progress, capital expenditures (separately for acquisitions and disposals). They also have to answer sections related to import or export of services and on the use of E-Commerce and employment, with further data on part-timers. However, the survey form can be sent in a long or in a short format. The main difference between the two types of formats is that in long format firms are required to provide a finer detail of the broad sections defined above. For instance, in the long format firms break down their disposals and acquisitions information about 20 different items, whereas in the short format they only report the aggregate values. Also, in the long format, firms answer on questions such as the total number of sites and the amount of squared metres they consist of.

2.3.1 Reporting units, selected and non-selected data

The ABI is covered by the Statistics of Trade Act, 1947, and therefore the firms are obliged by law to provide data if they get a form.[7] To reduce compliance costs however, the ABI is

[6] The ABI replaces Annual Employment Survey, Annual Census of Production and Construction (ACOP/ACOC), and the six following Annual Inquiries: wholesale, retail, motor trades, catering, property, service trades. In Catering and Allied Trades, between 1960 and 1979 there was a benchmark inquiry into catering roughly every four years or so, but from 1979 the inquiry became annual. There has been a property inquiry since the mid 1950s but until 1994 data was only collected on capital expenditure. From 1995, the range of data was extended to bring the inquiry in line with the other DS inquiries. Wholesaling and Dealing - The first major inquiry into wholesaling and dealing was carried out in respect of 1950, as part of the Census of Distribution. Subsequently, periodic large-scale detailed inquiries were conducted in respect of 1959, 1965, 1974 and 1990, but simpler annual inquiries were conducted for most intervening years and for all years since 1991. The first major inquiry into motor trades was carried out in 1950 as part of the Census of Distribution. Subsequently, periodic large-scale inquiries were conducted in respect of 1962, 1967 and 1972 although simple annual inquiries were carried out in most intervening years. By 1977 the annual inquiry was collecting detailed information on turnover and purchases. Regarding retailing, from 1950 periodic Censuses of Distribution were conducted, the last of which was in 1971. Full-scale inquiries covering every retail business and every retail outlet were taken for 1950, 1961 and 1971, with large-scale inquiries for 1957 and 1966. The first annual retailing inquiry was conducted in respect of 1976 with a sample of 30,000 units. Throughout the late 1970s and '80s the inquiry varied from year to year in terms of both sample size and the amount of information collected. From 1991-1997 the sample remained reasonably constant at around 12,000 (source for all this information: (http://www.statistics.gov.uk/StatBase/Source.asp?vlnk=1481&More=Y).

[7] Companies who have to fill out a form can refer to http://www.statistics.gov.uk/about/business_surveys/abi/default.asp for help and information.

not a Census of all local units. This is in two regards: aggregation and partial sampling. Regarding aggregation, enterprises normally report on all their local units jointly. There are two major exceptions. First, if the enterprise has local units in both Britain and Northern Ireland, there is a legal requirement for the ONS to keep data for these two areas separately and therefore enterprises are required to report data separately in this case. Second, there is separate reporting on LUs if a business explicitly requests such a split. So for example, Smith may decide to report on North and South combined and East and West separately. Thus returned data is at what is called the "reporting unit" (RU) level. The structure of RUs is shown for our example at the bottom of Figure 1. Brown forms one RU (A) only whereas Smith has two RUs (comprising of Smith North and Smith South; and Smith East and Smith West). Jones has one RU, comprising Jones North and Jones R&D.[8] Thus these RUs are the fundamental unit for reported data on the ARD.

It is worth noting at this point that the RU and LU distinction is crucial for our analysis. For example entry and exit at the LU level might look very different to that at the RU level. Regional issues are also important here: looking at RU data when an RU reports on a number of LUs where the LUs are based in different regions may give a very different picture to looking at LUs.

Regarding sampling, to reduce costs, only reporting units above a certain employment threshold (currently 250[9]) are all sent an ABI form every year. Smaller reporting units are sampled by size-region-industry bands.[10] In the ARD, all data returned from reporting units is held on what is called the "selected" file. Other data is held on the "non-selected file". Since the non-selected RUs are not sent a form, the non-selected data is of course the IDBR data.

In what follows we start by showing the number of and relation between LUs and RUs. We then ask what the data can be sensibly used for in analysis of size, entry/exit and productivity.

2.4 Evidence on RUs, LUs in retailing.

Let us now look at the numbers of LUs, RUs, etc. Table 2 sets out some of the relevant data. First, looking down column 1, top panel, there were 204,091 RUs in all retailing in 2001 and 292,115 LUs. Recall that RUs can report on one or more LUs, so the higher number of LUs is to be expected. Many of these RUs and LUs, by number, are in "other retail", "food, beverages, tobacco" and "non-specialised". The remainder of the top panel shows data on the

[8] On other surveys the RU structure might be slightly different, for example on the R&D survey Jones might report on Jones R&D only which would be its RU for that survey. This matters when matching surveys.

[9] The threshold was lower in the past. See Barnes and Martin (2002) for more details.

[10] The employment size bands are 1-9, 10-19, 20-49, 50-99, 100-249, the regions are England and Wales combined, Scotland and Northern Ireland. Within England and Wales industries are stratified at 4 digit level, NI is at two digit level and Scotland is at a hybrid 2/3/4 digit level (oversampling in Scotland and NI is by arrangement with local executives). See Partington (2001).

numbers of LUs that RUs report on. Column 3 shows that 11,527 RUs report on more than one LU. Thus, as column 4 shows, 192, 564 RUs, the bulk of the LUs, just report on one LU. The remaining columns sum up to 11,527 in column 3. So for example, the final column shows that only 160 RUs report on more than 100 LUs. In sum, approximately two-thirds of retailing outlets were accounted for by stand-alone businesses (192,564/292,115). Looking at the individual sectors, the distribution of units is the same in all seven.

These data are just numbers of RUs and LUs. The lower panel shows the average employment that these units account for. Here the picture, not surprisingly, is rather different. Columns 1 and 2 of the lower panel show mean employment in RU and LU (headcount, not FTE) is 13 and 9 in all retailing respectively. In the single LU RUs, it is 3.66. But looking at the last column, the RU who reports on more than 100 LUs has average employment per RU of almost 9,000. This figure, particularly, large in non-specialised, hints are concentration of employment, as we shall see later.

Table 2 suggests there are many LUs and RUs by number and considerable concentration of employment.

Table 3 gives some more details. Consider the top left panel, which shows data for all industries. The first number, 192, 564 is the same as in Table 2, column 4, top cell, namely the number of RUs who are stand-alone (i.e. responsible for only one LU). As the second column shows, this group accounts for 94.4% of the total number of RUs. Reading further across the table, total employment in these LUs is however 704, 769, which accounts for 26.5% of all employment. By contrast, looking at the bottom row of the top left panel, those reporting on more than 100 local units (160 RUs, 0.1% of total RUs), account for 54% of employment in all retailing. For non-specialised stores (including supermarkets), 75% of employment is accounted for by just 36 RUs, who are below 1% of the total number of RUs. Likewise in pharmaceuticals and other, the largest group accounts for under 2% of RUs by number., but 40% and 43% of total employment. By contrast second-hand stores are concentrated by both number and size in small groups., and so is, to a lesser extend, food, beverages and tobacco.[11]

The concentration of employment is shown in Table 4 which shows the percentage of the sector's employment in the top 5 and 10 RUs and LUs. Looking at the RU data, in non-specialised and not-in-stores, just ten stores account for over half of total employment. Looking at the LU data, the numbers are much less, since that is data by shop. This turns out to be an issue in the computation of productivity decompositions, see below.

[11] One issue for us is whether significant RUs change industry over time e.g. from many retailers are wholesalers as well and could be classified in different industries over time. To check this, we looked at the 6 largest supermarkets in the data set and found that they were consistently classified to one industry (SIC 52119). Evidently then, we do not have this problem in the data set for these companies.

Finally, the above data has shown the relation between RUs and LUs. Table 5 shows the relation between RUs and Enterprise groups. Most enterprise groups consist of one RU i.e. the mean number of RUs that each Enterprise group consists of is 1.01 in all sectors.

Thus far we have looked at employment by sector. What of regions? Table 6 shows mean employment by region. Consider first the average employment per RU in column 1. This is 20 in the Southeast, larger than elsewhere. The issue here of course is that RUs might report on a number of LUs and so if the RU is the head office, located in London, for example, this might be a misleading number for the average size of the actual store. Thus column 2 shows average size by LU. Here London is again the largest, but proportionately much less large than when measured by LU. To see whether these are significantly different, Table 7 shows the results of two analysis of variance exercises with employment, measured by LU and RU conditioned on year, region, industry and region*industry. In all cases LU and RU employment varies significantly across industries and regions and their interaction.

2.5 Section summary

We find

(a) In 2001 there were 292,115 stores in UK retailing and 204,091 firms/chains. But just 160 chains accounted for 54% of total employment

(b) Average store sizes are significantly different across regions and varies between 12 employees in the South East and 4 in Wales.

3 Entry and exit

This section looks at exit and entry defined as

TYPE	DEFINITION
Entrant	Present in t and not present in t – 1
Exitor	Present in t and not present in t + 1
1-year	Present in t and not present in either t – 1 or t + 1
Stayer	Neither of the above three

We look separately at RUs and LUs to provide as full as information as possible. Using these definitions, the basic data for the whole retailing sector covering the period 1998-2001 is set out in Table 8. The total numbers of RUs and LUs are as shown in the total column, and the numbers in the left-hand four columns add up to this number. As it shows, the bulk of the RUs and LUs are stayers with entry and exit rates (i.e. entry and exit numbers as shares of the total number of LUs that year) of around 10-20% depending a bit on RU or LU status. Note the apparently high entry rate in 1998 by LU and RU, which likely reflects underlying problems with the business register in 1997.

Table 9 and Table 10 show data on entry and exit rates by industry and region. By industry, entry and exit rates for LUs (lower panel of Table 9) look quite similar, with 14% in "Not in Stores" and 6.79% in "Food, beverages and tobacco" being the maximum and minimum entry rates and 16.09 and 10.99 being likewise the maximum and minimum exit rates. Regional differences in entry and exit rates by LUs (lower panel of Table 10) are quite small with maxima and minima of 10.84 and 8.20 and 13.58 and 11.20 respectively.

Since it has been argued that restrictive zoning policy has affected UK productivity it is of interest to see if these numbers are significantly different across regions, assuming that zoning policy differs across regions. Thus we looked at a similar analysis of variance by analysing regional entry and exit, measured by LU by region, industry and region times industry, see Table 11. We conclude from this table that industry – rather than regional - differences explain more of the variation in the data.

We now consider whether there are differences in employment. Table 12 shows data on mean employment (measured by LU) by entry, exit, stayer and one-year status for regions and industries for 2001 (these data are quite stable over time). Looking at industry first, employment of stayers is larger in "non-specialised", at about 20 and smallest in second-hand, at 2.66. Entrants and exitors tend to be similar sized except in non-specialised where entrants are rather larger then exitors and stayers. The one-year LUs are sometimes quite

large.[12] Turning to the regional figures, the sizes of are rather similar across regions with the exception of the South-East where all types – except exitors -seem to be larger.

3.1 Section summary

We find:

(a) Entry/exit/stayers are stable fractions of all stores being about 19%, 8% and 69% (the rest are stores who survive one year).

(b) Entry and exit rates are highest in "Not in Stores" and "Food, tobacco and beverages" respectively and in the South East (and Scotland) and North West respectively.

(c) By employment, entrants and stayers are largest in the South East.

4 Productivity

4.1 What productivity data is available?

As discussed above, data is at essentially two levels, RU and LU. RU data is returned data i.e. it relies on data actually reported by firms. LU data is a mix of data that is from the ARI, and so is reported by firms, and from other sources e.g. taxes, which is inferred. Now, given that the LUs correspond to stores, this would seem to be the most desirable for a number of cases, especially since a number of retailers consist of many stores. Thus can we use the LU data? For single unit RUs we can use the RU data which is the LU data. Thus the question is whether we can use the LU data that does not correspond to single unit RUs. That data is from the IDBR data, which is either from the IDBR administrative sources i.e. the VAT or PAYE, or other data that brought the business onto the register in the first place, or the ARI. The following points regarding these data are worth noting.

First, since some of the input data is interpolated from sales data and vice versa, one cannot do productivity studies. Second, there is a specific problem with employment data on the IDBR. According to ONS (2001), when a business first arrives on the register, its employment, if present, is frozen at its first reported point until updated. Turnover is updated however. Thus productivity for these businesses is unreliable unless their employment is updated. Updating is done from the results of the ARI, or before the ARI was introduced, if the firm was in one of the Annual Employment Surveys (AES). We can get some impression of the problem by considering Table 13. The table shows when the employment data of enterprises in the IDBR in year 2000 were last updated. The first 4 columns of Table 4 refer to different size bands. The final column shows that 8.5% of total employment had not been

[12] One concern here is that LUs who change industry, perhaps due to misreporting, might show up as entrants and exitors. To prevent this from distorting the numbers, for LUs who changed industry only once, we assigned LUs to their modal industry throughout the period.

updated since 1993. 1993 is the year when there was last a Census of Employment. Looking at columns 1 and 2 we see that the updating problem is concentrated in the smallest enterprises. 28.7% of employment in enterprises of size 0-9 and 40.2% of employment in enterprises of size 10-19 had not been updated since 1993. Indeed row 11 of Table 4 also reports that 56.9 and 21.8 of enterprises of size 0-9 and 10-19 have *never* been sent an ARI form or included in the AES. By contrast, larger enterprises are updated more frequently. An additional problem is that the ONS (2001) also state that even larger enterprises in the ARI or AES, may not have fully reported on their local units (see also Partington, 2001)[13]. This suggests that the employment and therefore productivity data on these smaller enterprises, who are overwhelmingly in the non-selected sample, is likely very unreliable.
In what follows therefore, we use productivity data at the reporting unit level, from the selected file.

4.2 Data available on outputs and inputs

As described above, the only reliable input and output data is that available for RUs and Table 14 sets out some of the basic data available for all retailing sectors. Each observation in the data represents one RU. The top rows show data on sales, gross value added and gross output. Following the ONS, gross value added at factor cost is calculated as follows:

Gross value added at factor cost =Turnover (exc. VAT) + ? Total stocks (*dstocks*)
+ Work of Capital Nature by Own Staff (*capaq*) + Insurance Claims Received (*ins*)
- Purchases (*purch*)

Gross output, on the other hand, is:

Gross output= Turnover (exc. VAT)+ ? Work in Progress+ ? Stocks Bought for Resale
+ Work of Capital Nature by Own Staff.

The main difference between the two being the purchases figure, which is deducted in the calculation of gross value added. The rest of the table above shows some summary statistics for each variable; not suprisingly, purchases being the largest element after sales.
One interesting point is that we have data on employment and the fraction of employees who are part-time. We do not know however what proportion of the full week such workers work, so we allocated them to 50% of the work week to calculate FTEs. Thus in what follows we present productivity data by employment and by FTE employment.

[13] Partington (2001) states that the AES sent *x* LU forms to each multi-LU enterprise with *x* based on the expected number of LUs according to administrative sources. Enterprises with less LUs disposed of excess forms, but since there was no systematic method of obtaining more forms, RUs with more LUs than expected simply did not report on these "excess" LUs.

4.3 Deflators

Published data on services produced by the ONS uses different deflators depending on the nature of the data being presented. The most commonly used are GDP deflators, but others include the Retail Price Index, and the Producer Price Index. For the retail sector, the ONS produces a separate index called the Retail Sales Index, which collects retail sales figures on a monthly basis. This is used to produce a disaggregated price index for the 4 digit SIC codes within the retailing sector, and is therefore covering in detail the specific area we are looking at. We believe this is the most appropriate index to use for our particular work, due to its concern with only those firms with retail SIC codes, and due to how it is measured.

4.4 Productivity in retailing: Definitions.

What is the correct measure of productivity in retailing? A number of papers discuss this issue. Triplett and Bosworth (2002) note that the BLS uses turnover per employee and the BEA gross margin (sales less goods sold). We can also construct gross margins, but we then have the problem of double deflation since there are no input deflators for retailing. Thus, for example, the shift of sales from a high price department store to a low price discount store is treated as a reduction in quality not a reduction in price. One method is therefore to try to deflate margins directly by price indices that account for some of the variation in margins such as floor area, stock keeping units, type of store (discount etc). Since we do not have these data, for output we use gross output and value added. For employment we use employment and FTE employment..[14]

4.5 Weights

Since we use the selected file we deal not with the whole industry, but a sample. Thus we need to develop weights to use where appropriate. To do this we use both the selected and non-selected file, but with a robustness check as follows. We combined both files to make a grand file of selected RUs and non-selected LUs. We then split the sample into six size bands (0-9, 10-20, 21-50, 51-100, 101-250 and 250+). We then calculated weights as the sum of selected and non-selected employment divided by selected employment in each sizeband. So for example, if a firm falls into sizeband 50-99 and total selected employment was 1,000 but total selected and non-selected employment was 2,000, the weight for that sizeband would be 2. We checked to see that no weight was abnormally large.

[14] In their US study, Foster, Haltiwanger and Krizan (2001) calculate productivity decompositions for US retailing using Census data for 1987, 1992 and 1997. Productivity is sales per employee, with employees multiplied by BLS average industry hours to adjust for hours, and sales deflated by four-digit deflators. They have data for establishments, which are the equivalent of local units.

4.6 Productivity findings

Table 15 sets out data, by year, for all sectors data on productivity measured with output as gross output and gross value added and employment as headcount and FTE, with the total number of observations in listed too. A number of points are worth noting. First, the 1997 data appear to be slightly greater than the 1998 data, suggesting negative labour productivity growth. This could be due to problems with the 1997 data. Second, there is interesting variation by per person and per FTE. The levels are greater with FTEs in both cases, as would be expected, but the growth rates are similar.

Third, it is of interest to check these data against official data, noting that the ARD is a research tool and may not match official data sources. The official data available to use is set out in the lower part of the table, under the heading *"published"*. The first row shows the growth rate (of value added per person, 1998-2002) computed from the published ABI tables. These tables are the aggregated numbers published by the ONS based on the disaggregated ABI data we have used. These match our numbers very closely which is reassuring given the uncertainties over weighting.

Whilst these ABI numbers are released, they are not the official productivity growth figures, which are instead provided for more aggregated industries and to be congruent with the National Accounts. The next two rows report these. They are first the official published experimental data for labour productivity for the non-production industries, based on Daffin, Reed and Vaze (2002). These data are not available for retailing alone, but for retailing, distribution, hotels and catering. The numbers are much lower than ours, but this could reflect low productivity growth in the other sectors. Finally, the final row shows the comparable numbers from the EU KLEMS project data which builds on the OECD STAN project for the UK and, for comparison, the US. Again, our UK numbers are rather higher than these.

4.7 Productivity levels by size

Table 16 looks at productivity levels by size of RU. The top panel shows all sectors with productivity calculated by employment and by FTE, and the lower panel shows the non-specialised industry. As the table shows, all productivity levels rise when one compares headcounts and FTEs. But it is interesting that the smaller firms productivity seems to rise much more. Using employment for example, it would seem that the small firm's productivity advantage in all sectors is 24%=(23.59/19.06-1), whereas using FTEs it is 42% (=32.86/23.21-1). In supermarkets the larger firms have a productivity advantage over other classes except for the very smallest firms, but it is less of an advantage when using FTEs.

4.8 Productivity spreads

Table 17 contains data on productivity spreads. Foster et al (2002) quote a standard deviation and interquartile range of 0.5 for hours-weighted log gross output per head in after taking deviations from four-digit means. We use data on log gross output per FTE in after taking deviations from three-digit means. As the table shows, we find a slightly higher standard deviation and interquartile range than they do.

5 The sources of productivity growth

What is the contribution of entry and exit to productivity growth in services? We employ the decomposition of Foster, Haltiwanger, and Krizan (FHK, 1998). Start by writing manufacturing-wide productivity in year t, P_t as:

$$P_t = \sum_i q_{it} p_{it} \qquad (1)$$

where θ_i is the share of establishment i (employment share) and p_{it} is $\ln(GVA_i/n_i)$. Foster, Haltiwanger and Krizan (1998) (FHK) suggest a decomposition to the change in manufacturing-wide labour productivity or lnTFP between t-k and t, ΔP_t as

$$\Delta P_t = \sum_{i\in S} q_{i,t-k}\Delta p_{it} + \sum_{i\in S}\Delta q_{it}(p_{i,t-k} - P_{t-k}) + \sum_{i\in S}\Delta q_{it}\Delta p_{it} + \sum_{i\in N} q_{it}(p_{it} - P_{t-k}) - \sum_{i\in X} q_{i,t-k}(p_{i,t-k} - P_{t-k}) \qquad \text{FHK} \quad (2)$$

where S, N and X denotes the establishments that survive, enter and exit respectively between t and t-k. We denote this the FHK decomposition. The first term in (2) shows the contribution to productivity growth of growth within the surviving establishments, or the "within" effect. The second term shows the contribution of changes in shares of the survivors weighted by start period productivity relative to the average (often termed the "between" effect). So, productivity grows if the shares of higher productivity establishments increase. The third term is an additional covariance term that is positive when market share increases (falls) for establishments with growing (falling) productivity. The fourth and fifth terms show the contribution of entry and exit.[15] They are positive when there is entry (exit) of above- (below-) average productivity establishments.

To calculate this we proceed as follows. First, we performed the decomposition 1998-2002. Second, recall that RUs can exit and enter from the selected file either if they truly have

[15] With industry data one can decompose DP_t into the within and between terms, but cannot of course account for net entry. See Cameron et al (1998) for implementation of this on UK data.

entered and exited or if they have moved to the non-selected file. Thus we use the selected and non-selected data to identify true exitors and entrants. Third, we dropped all non-selected firms, since we do not have productivity information for them, and all those firms who falsely entered and exited the selected sample; we do not have productivity information for them in at least one year and we do not want to ascribe a part of entry and exit to them. Therefore, we therefore identify survivorship, entry and exit for a set of firms who are selected and do not falsely enter and exit the selected file. Fourth, we calculate two sets of weights: employment weights and also employment weights taking into account sampling.[16]

Fifth, we perform these calculations by three-digit industry, i.e. the P in (2) is the three-digit average industry productivity level and the θ is the share of three-digit industry employment. Sixth, all employment data are FTE employment. Seventh, the overall numbers for all industries are constructed as a weighted sum of the numbers for the individual industries, where, following Foster et al (2002) the weights are the share of gross value added (since we use value added as our productivity measure in the decomposition) in each industry averaged over the start and end period. Seventh, the data are deflated by prices from the Retail Sales Inquiry values as above.

The results are shown in Table 18. This uses gross value added per FTE as the productivity measure. The top panel uses all RUs and the lower drops some, see below. Consider first the top panel. This shows, for each industry, the contribution of stayers and the sum of entry and exitors. The left hand panel performs the decomposition without sampling weights and the right with sampling weights. The italicised numbers on the right are the (gross output) weights of each industry in overall gross output and the frequencies of each type.

The top row is somewhat complicated so consider the second row, for "food, beverages and tobacco" first. Overall productivity growth is 7% over the whole 4 years. The contribution of stayers (49 RUs) and entry/exit (341 and 170 RUs) is over one, meaning that each group contributed more than overall productivity growth. In this case entry and exit was in fact a drag on productivity growth. Looking at the other industries, the effect of entry and exit is beneficial in all cases except in "not in stores".

Turning to the top row, i.e. the non-specialised sector, which as 104 stayers, 287 exitors and 203 entrants. Productivity growth is negative. In the raw data, the effect of entry and exit is positive and that of stayers negative. Thus in the table we write the entry/exit term as being positive, since the entry/exit process helped productivity growth and the stayer term as negative since it hurt productivity growth. We do this in all cases where productivity growth is negative i.e. the term in the cell is the effect as a fraction of overall productivity growth but

[16] The former are straightforward being employment for unit i divided by employment in all i units in the industry. The latter is employment in plant i times the weight that plant has, divided b the sum of thus weighted employment in the industry.

the sign takes the sign of the contribution.[17] Thus in all cases for "non-specialised" stayers had a negative impact on productivity growth and entry/exit a positive effect. Note that some of these numbers are very large because the underlying (4 year) productivity growth figures are rather small.

The overall picture is that the entry/exit process contributes positively to productivity growth accounting for 70% of overall productivity growth with no sampling weights and 82% with sampling weights.

There are at least two robustness checks of these numbers that suggest themselves. First, as the "ind weight" column indicates, much of retailing is in the "non-specialised" category. If we recompute the overall figure leaving this category out (and reweighting accordingly) we obtain an overall contribution of stayers of 40% and of entry/exit of 60%.

Second, as we have seen, when using LU data, each LU has a small share of total employment. However, since sectors of retailing are so concentrated, the largest RUs potentially have a very large θ (think of a Wal-Mart store versus all of Wal-Mart). Thus these calculations might be sensitive to whether or not one includes a large store, if, for example, that large store exited, or had particularly poor productivity growth. In a sense since the decompositions add up the overall productivity growth one is simply decomposing a different summation of productivity growth, but nonetheless one might reasonably argue that it is important to see if the overall conclusions on entry and exit are influenced. Thus the bottom panel of Table 18 examines this by dropping the top two RUs, by weight size, in each industry. As the panel shows the overall entry/exit contribution falls but is still more than half the total.

How do these numbers compare with the US data, which shows almost 100% of productivity growth between 1987-97 (and subperiods) to be due to entry and exit?[18] The US data are for individual stores, not firms as we have here. However, Table 3 of FHK also provides data on the fraction of entry and exit due to expansion and closure of stores within existing firms. This shows that 40% of all productivity growth is due to this source. Thus a US decomposition using firms would show that 40% of productivity growth is due to within-firm effects and 60% due to entry and exit. This is not that far away from what we find here.

Finally, we wish also to look at the decompositions by region. However, there are two problems with this. First, we have very few regional observations to do this by industry and were uncomfortable with analysing the contributions of a region consisting of many industries relative to average productivity of the region. Second, as discussed extensively above, regional entry and exit of RUs is not likely to be the relevant concept for seeing if planning restrictions, which presumably impact on LUs, matter.

---

[17] This does not affect the overall share numbers since they are calculated as a share-weighted stayer and entry/exit numbers divided by share-weighted total productivity growth.

[18] Our data is a shorter subperiod than the 5 years that FHK use, but the subperiods they use still show the same fraction of productivity growth due to entry and exit in the longer period.

5.1 Section summary

We find

(a) Productivity is best measured at the RU level

(b) Aggregate productivity growth is higher than that published

(c) The variation in labour productivity across retailers is rather larger than in the US

(d) If anything, the contribution of entry and exit to productivity growth is somewhat larger than in the US.

5.2 Have restrictive zoning laws held UK productivity back?

On the basis of these numbers what can we say about the proposition that zoning laws have hindered productivity growth in UK retailing? First, it depends what effect zoning has. Suppose first that there are economies of scale in retailing and regulations stop firms from entering at MES. Thus it likely affects the entry of large shops and means that there is less possible avenue for entry to boost productivity growth. On the face of it this would seem not to be supported by the data, for we find comparable entry and exit contributions to the US, if not a bit higher. One caveat is that since our data are at the RU level, entry and exit is dominated by RUs accounted for by single-LU RUs, which, as we have seen are small anyway. But if that was the case, and such entrants are suboptimally small, then we would see even less effect from entry and exit.

Another possibility is that zoning laws make land and entry expensive so that only high productivity businesses can enter. If this is the case we might expect a greater contribution from entry and exit. We find this. It would be best to investigate more of this entry and exit contribution to better establish this point: it would, for example, be of great value to compare the average entry/exit advantage/disadvantage relative to the stayers in different countries.

Another avenue for future work is to see if regulation differs across regions and whether this is reflect in different entry and exit effects. The RU problem is potentially acute here, since regulation would likely affect LUs but have ambiguous effect on multi-LU RUs. We could however investigate the average entry rates of LUs, but not their contribution to productivity growth.

6 Conclusions

We have used a new micro-level data set to study productivity in UK retailing, 1997-2002. We have used store level data to look at concentration and entry and exit, but, due to data limitations, chain of stores level data to look at productivity and productivity growth. Among our findings are:

(a) In 2001 there were 292,115 stores in UK retailing and 204,091 firms/chains. But just 160 chains accounted for 54% of total employment

(b) Average store sizes, entry and exit rates are significantly different across regions and industries

(c) Entry/exit/stayers are stable fractions of all stores being about 13%, 13% and 69% (the rest are stores who survive one year).

(d) Productivity levels are strongly affected by whether productivity is measured by heads or full-time equivalents.

(e) Implied aggregate productivity growth from our data is somewhat higher than that published

(f) The variation in labour productivity across retailers is rather larger than in the US

(g) If anything, the contribution of entry and exit to productivity growth is somewhat larger than in the US.

Figure 1: Plants and firms in the IDBR

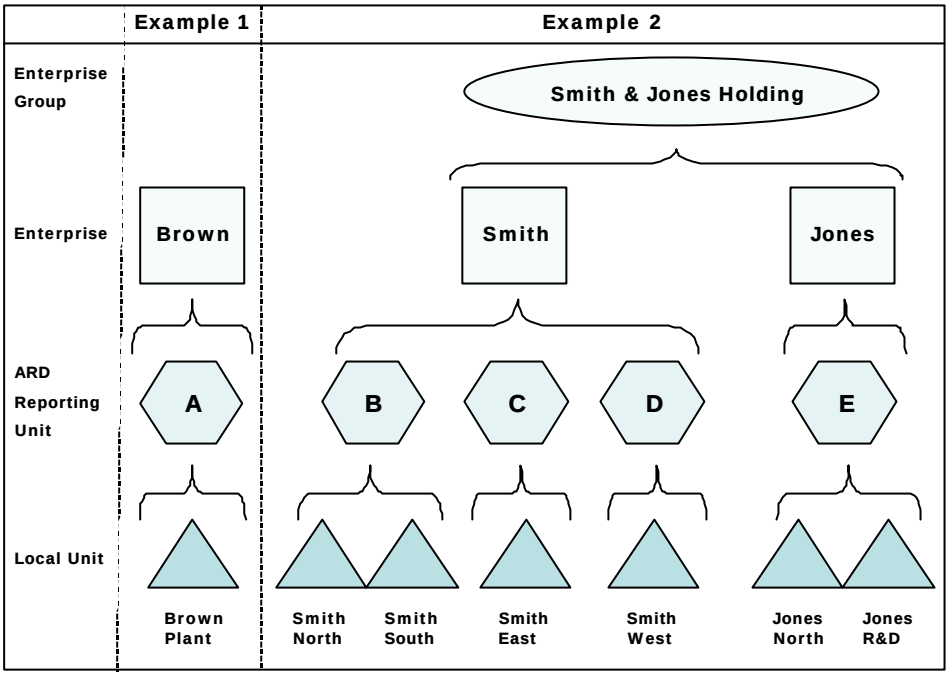

Table 1: Industries covered in UK ARD retailing data

SIC code	Industry	Notes
521	Retail sales in Non-spec covering e.g. food, beverages or tobacco	Includes supermarkets and department stores
522	Food, beverages, tobacco in specialised stores	
523	Pharm and medical goods, cosmetic and toilet articles	Includes chemists
524	Other retail sales of new goods in specialised stores	Includes sales of textiles, clothing, shoes, furniture, elect appliances, hardware, books, newspapers and stationary, cameras, office supplies, computers. Clothing is biggest area
525	Second-hand	Mostly second-hand books, second-hand goods and antiques
526	Not in stores	Mostly mail order and stalls and markets
527	Repair	Repair of personal goods, boots and shoes, watches and clocks

Table 2: Reporting Unit and Local Unit Numbers

Year=2001	1	2	3	4	5	6	7	8
			Frequencies					
Sectors	**Total # of RU**	**Total # of LU**	**# of RU with more than 1 LU**	**RU with 1 LU**	**RU with 2 to 5 LU**	**RU with 6 to 10 LU**	**RU with 11 to 100 LU**	**RU with more than 100 LU**
52. All Sectors	204,091	292,115	11,527	192,564	10,179	644	544	160
521. Non-specialised	36,615	55,673	998	35,617	823	69	70	36
522. Food, beverages, tobacco	40,027	50,924	1,873	38,154	1,675	111	<100	<20
523. Pharmaceutical	6,560	12,399	836	5,724	734	<60	<50	<10
524. Other retail	94,679	142,814	7,224	87,455	6,409	382	337	96
525. Second-hand	6,035	7,373	152	5,883	<150	<10	<10	<10
526. Not in stores	14,211	15,470	301	13,910	267	<20	<20	<20
527. Repair	5,964	7,462	143	5,821	<150	<10	<10	<10
			Mean Employment					
Sectors	**Mean RU Employment**	**Mean LU Employment**		**RU with 1 LU**	**RU with 2 to 5 LU**	**RU with 6 to 10 LU**	**RU with 11 to 100 LU**	**RU with more than 100 LU**
52. All Sectors	13.03	9.11		3.66	16.93	99.67	516.83	8,986.54
521. Non-specialised	31.40	21.02		3.67	35.76	239.81	1,522.23	24,075.20
522. Food, beverages, tobacco	5.44	4.21		3.26	15.05	51.62	295.60	2,895.71
523. Pharmaceutical	13.04	6.95		5.07	15.35	56.22	<200	<5000
524. Other retail	10.97	7.09		3.92	13.79	70.57	<400	<5000
525. Second-hand	3.24	2.58		2.55	9.38	17.14	66.17	920.00
526. Not in stores	8.53	6.65		3.18	<60	<700	<700	<21000
527. Repair	4.65	3.45		3.17	10.96	<50	<1000	<2000

Note: some of the cells have been suppressed for disclosure reasons. Columns 4, 5, 6, 7 and 8 add up to column 1. Columns 5,6,7 and 8 add to column 3.
Source: authors' calculations from ARD.

Table 3 Employment in largest firms, by sector, 2001

Year=2001	1	2	3	4	5	6		7	8	9	10	11	12
Sector	# of LU belonging to RU	Freq	% of Total # of RU	Total Emp	% of Total Emp	Mean Empl	Sector	# of LU belonging to RU	Freq	% of Total # of RU	Total Emp	% of Total Emp	Mean Empl
52. All Sectors	0-	192,564	94.4	704,769	26.5	3.7	**524. Other retail**	0-	87,455	92.4	342,939	33.0	3.9
	2-	10,179	5.0	172,313	6.5	16.9		2-	6,409	6.8	88,396	8.5	13.8
	6-	644	0.3	64,185	2.4	99.7		6-	382	0.4	26,958	2.6	70.6
	11-	544	0.3	281,155	10.6	516.8		11-	337	0.4	132,580	12.8	393.4
	101-	160	0.1	1,437,846	54.0	8,986.5		101-	96	0.1	447,872	43.1	4,665.3
521. Non-specialised	0-	35,617	97.3	130,635	11.4	3.7	**525. Second-hand**	0-	5,883	97.5	15,011	76.7	2.6
	2-	823	2.2	29,431		35.8		2-	<150	2.3	1,276		9.4
	6-	69		16,547		239.8		6-	<10		120		17.1
	11-	70		106,556		1,522.2		11-	<10		397		66.2
	101-	36		866,707	75.4	24,075.2		101-	<10		2,760	14.1	920.0
522. Food, beverages, tobacco	0-	38,154	95.3	124,566	57.2	3.3	**526. Not in stores**	0-	13,910	97.9	44,185	36.5	3.2
	2-	1,675	4.2	25,202		15.0		2-	267	1.9	15,259	12.6	57.1
	6-	111	0.3	5,730		51.6		6-	<20		11,694		615.5
	11-	<100		21,579		295.6		11-	<20		8,736		672.0
	101-	<20		40,540	18.6	2,895.7		101-	<10		41,333		20,666.5
523. Pharmaceutical	0-	5,724	87.3	29,001	33.9	5.1	**527. Repair**	0-	5,821	97.6	18,432	66.5	3.2
	2-	734	11.2	11,270		15.4		2-	135		1,479		11.0
	6-	<60		3,036		56.2		6-	<10		100		50.0
	11-	<50		7,573		184.7		11-	<10		3,734		933.5
	101-	<10		34,674	40.5	4,953.4		101-	<10		3,960		1,980.0

Note: Some of the cells have been suppressed for disclosures reasons.

Source: authors' calculations from ARD

Table 4: Concentration of Employment by industry, five and ten firm concentration ratios, by RU and LU

Year=2001

Reporting Units

Industry	cr5	cr10	Number of RU
52. All Sectors	20.53	28.23	204,091
521. Non-specialised	47.29	62.17	36,615
522. Food, beverages, tobacco	13.02	17.06	40,027
523. Pharmaceutical	38.16	42.64	6,560
524. Other retail	15.28	22.10	94,679
525. Second-hand	15.85	18.08	6,035
526. Not in stores	44.37	50.97	14,211
527. Repair	30.24	32.42	5,964

Local Units

Industry	cr5	cr10	Number of LU
52. All Sectors	0.54	0.95	292,122
521. Non-specialised	1.16	1.90	55,674
522. Food, beverages, tobacco	0.74	1.18	50,928
523. Pharmaceutical	1.58	2.36	12,399
524. Other retail	0.91	1.44	142,815
525. Second-hand	3.72	5.32	7,374
526. Not in stores	8.67	14.93	15,470
527. Repair	6.85	9.51	7,462

Source: authors' calculations from ARD.

Table 5: Relation between reporting units and enterprise groups

Reporting Unit per Enterprise Group

Industry	Mean	Standard Deviation
52. All Sectors	1.01	0.24
521. Non-specialised	1.00	0.25
522. Food, beverages, tobacco	1.00	0.08
523. Pharmaceutical	1.02	0.20
524. Other retail	1.01	0.29
525. Second-hand	1.00	0.05
526. Not in stores	1.00	0.09
527. Repair	1.00	0.06
52. All Sectors	1.44	14.66

Source: authors' calculations from ARD.

Table 6: Mean employment, by Region

Year=2001, All Industries

Region	Employment per Reporting Unit	Employment per Local Unit	Frequency
South East (G)	20.32	12.09	67992
East Anglia (F)	6.56	5.61	7566
South West (J)	9.74	7.52	18787
West Midlands (E)	6.11	4.82	18057
East Midlands (C)	6.94	5.55	14326
Yorkshire & Humberside (C)	9.23	6.92	17963
North West (B)	13.66	9.89	23014
North (A)	8.86	6.65	9274
Wales (W)	5.31	4.47	9935
Scotland (x)	12.88	9.22	17177

Source: authors' calculations from ARD.

Table 7: Does mean employment differ across regions?

Anova. Mean Employment per Local Unit		
Variable	F Test	Prob > F
Year	4.07	0.00
Industry	2770.99	0.00
Region	21.82	0.00
Region*Industry	22.01	0.00
All	243.66	0.00
Adj R-square	0.98	

Anova. Mean Employment per Reporting Unit		
Variable	F Test	Prob > F
Year	0.19	0.97
Industry	364.04	0.00
Region	72.38	0.00
Region*Industry	40.58	0.00
All	67.95	0.00
Adj R-square	0.92	

Table 8: Entry & Exit

Reporting Units

	Frequencies					Rates			
Year	**Stayers**	**Entrants**	**Exitors**	**1 Year**	**Total**	**Entry Rate**	**Exit Rate**	**Stay Rate**	**1 Year Rate**
1998	144,174	39,923	18,391	6,879	209,367	19.07	8.78	68.86	3.29
1999	163,114	21,513	20,983	4,792	210,402	10.22	9.97	77.52	2.28
2000	163,546	20,192	21,081	4,461	209,280	9.65	10.07	78.15	2.13
2001	162,120	17,300	21,618	3,630	204,668	8.45	10.56	79.21	1.77

Local Units

	Frequencies					Rates			
Year	**Stayers**	**Entrants**	**Exitors**	**1 Year**	**Total**	**Entry Rate**	**Exit Rate**	**Stay Rate**	**1 Year Rate**
1998	200,723	50,074	25,833	8,900	285,530	17.54	9.05	70.30	3.12
1999	220,665	30,008	30,132	6,857	287,662	10.43	10.47	76.71	2.38
2000	213,371	28,369	37,302	7,474	286,516	9.90	13.02	74.47	2.61
2001	206,936	28,265	34,804	10,161	280,166	10.09	12.42	73.86	3.63

Source: authors' calculations from ARD.

Table 9: Entry and exit rates by industry

Reporting Units

Year=2001	Frequencies					Rates			
Sectors	**Stayers**	**Entrants**	**Exitors**	**1 Year**	**Total**	**Entry Rate**	**Exit Rate**	**Stay Rate**	**1 Year Rate**
52. All Sectors	162,120	17,300	21,618	3,630	204,668	8.45	10.56	79.21	1.77
521. Non-specialised	28,329	3,638	4,091	645	36,703	9.91	11.15	77.18	1.76
522. Food, beverages, tobacco	33,087	2,302	4,269	479	40,137	5.74	10.64	82.44	1.19
523. Pharmaceutical	5,470	433	612	67	6,582	6.58	9.30	83.11	1.02
524. Other retail	76,250	7,856	9,261	1,588	94,955	8.27	9.75	80.30	1.67
525. Second-hand	5,007	393	540	124	6,064	6.48	8.91	82.57	2.04
526. Not in stores	9,629	1,902	2,185	525	14,241	13.36	15.34	67.61	3.69
527. Repair	4,348	776	660	202	5,986	12.96	11.03	72.64	3.37

Local Units

Year=2001	Frequencies					Rates			
Sectors	**Stayers**	**Entrants**	**Exitors**	**1 Year**	**Total**	**Entry Rate**	**Exit Rate**	**Stay Rate**	**1 Year Rate**
52. All Sectors	206,936	28,265	34,804	10,161	280,166	10.09	12.42	73.86	3.63
521. Non-specialised	36,508	5,939	7,012	2,799	52,258	11.36	13.42	69.86	5.36
522. Food, beverages, tobacco	37,169	3,304	7,377	824	48,674	6.79	15.16	76.36	1.69
523. Pharmaceutical	8,708	1,275	1,289	360	11,632	10.96	11.08	74.86	3.09
524. Other retail	101,275	13,851	14,783	4,634	134,543	10.29	10.99	75.27	3.44
525. Second-hand	5,417	598	745	340	7,100	8.42	10.49	76.30	4.79
526. Not in stores	9,735	2,040	2,374	608	14,757	13.82	16.09	65.97	4.12
527. Repair	4,587	855	737	469	6,648	12.86	11.09	69.00	7.05

Source: authors' calculations from ARD.

Table 10: Entry and exit rates by region

Reporting Units

Year=2001	**Frequencies**					**Rates**			
Regions	**Stayers**	**Entrants**	**Exitors**	**1 Year**	**Total**	**Entry Rate**	**Exit Rate**	**Stay Rate**	**1 Year Rate**
South East (G)	53,051	6,274	7,556	1,373	68,254	9.19	11.07	77.73	2.01
East Anglia (F)	6,158	582	745	102	7,587	7.67	9.82	81.17	1.34
South West (J)	14,915	1,630	1,991	299	18,835	8.65	10.57	79.19	1.59
West Midlands (E)	14,461	1,411	1,919	302	18,093	7.80	10.61	79.93	1.67
East Midlands (C)	11,578	1,105	1,428	243	14,354	7.70	9.95	80.66	1.69
Yorkshire & Humberside (C)	14,626	1,419	1,718	239	18,002	7.88	9.54	81.25	1.33
North West (B)	18,060	1,893	2,658	453	23,064	8.21	11.52	78.30	1.96
North (A)	7,669	668	843	111	9,291	7.19	9.07	82.54	1.19
Wales (W)	8,125	732	957	145	9,959	7.35	9.61	81.58	1.46
Scotland (x)	13,477	1,586	1,803	363	17,229	9.21	10.46	78.22	2.11

Local Units

Year=2001	**Frequencies**					**Rates**			
Regions	**Stayers**	**Entrants**	**Exitors**	**1 Year**	**Total**	**Entry Rate**	**Exit Rate**	**Stay Rate**	**1 Year Rate**
South East (G)	66,014	9,858	11,694	3,390	90,956	10.84	12.86	72.58	3.73
South West (J)	18,871	2,756	3,130	875	25,632	10.75	12.21	73.62	3.41
East Anglia (F)	7,841	1,019	1,155	298	10,313	9.88	11.20	76.03	2.89
West Midlands (E)	18,617	2,349	2,980	977	24,923	9.43	11.96	74.70	3.92
East Midlands (C)	14,565	1,821	2,252	650	19,288	9.44	11.68	75.51	3.37
Yorkshire & Humberside (C)	18,843	2,343	2,787	687	24,660	9.50	11.30	76.41	2.79
North West (B)	23,081	3,042	4,129	1,500	31,752	9.58	13.00	72.69	4.72
North (A)	10,572	1,121	1,596	388	13,677	8.20	11.67	77.30	2.84
Wales (W)	10,380	1,248	1,611	425	13,664	9.13	11.79	75.97	3.11
Scotland (x)	18,148	2,707	3,424	940	25,219	10.73	13.58	71.96	3.73

Source: authors' calculations from ARD.

Table 11: ANOVA of entry and exit rates

Anova. Entry Rates at the Local Unit Level		
Variable	F Test	Prob > F
Year	30.82	0.00
Industry	50.58	0.00
Region	1.23	0.28
Region*Industry	0.41	1.00
All	5.90	0.00
Adj R-square	0.56	

Anova. Entry Rates at the Reporting Unit Level		
Variable	F Test	Prob > F
Year	33.96	0.00
Industry	56.51	0.00
Region	1.13	0.35
Region*Industry	0.33	1.00
All	6.45	0.00
Adj R-square	0.58	

Anova. Exit Rates at the Local Unit Level		
Variable	F Test	Prob > F
Year	53.33	0.00
Industry	14.00	0.00
Region	1.14	0.34
Region*Industry	0.51	1.00
All	3.66	0.00
Adj R-square	0.41	

Anova. Exit Rates at the Reporting Unit Level		
Variable	F Test	Prob > F
Year	19.77	0.00
Industry	33.25	0.00
Region	3.16	0.00
Region*Industry	0.84	0.79
All	4.56	0.00
Adj R-square	0.48	

Table 12: Mean LU employment by entry, exit status, by region and industry

Local Units					
Industry	**Stayers**	**Entrants**	**Exitors**	**1 Year**	**All**
52. All Sectors	8.07	11.46	7.72	15.18	8.62
521. Non-specialised	17.22	32.54	13.41	33.65	19.33
522. Food, beverages, tobacco	4.16	3.63	4.19	4.66	4.14
523. Pharmaceutical	6.38	10.19	6.15	7.44	6.81
524. Other retail	6.69	6.47	7.33	9.64	6.84
525. Second-hand	2.66	2.14	2.31	2.49	2.57
526. Not in stores	7.02	3.41	6.33	5.18	6.33
527. Repair	3.42	2.47	3.59	4.81	3.42

Region	**Stayers**	**Entrants**	**Exitors**	**1 Year**	**All**
South East (G)	8.62	12.49	7.49	14.20	9.10
East Anglia (F)	8.29	11.52	7.82	11.61	8.65
South West (J)	7.83	12.97	7.75	12.11	8.52
West Midlands (E)	7.51	11.92	5.87	15.77	8.05
East Midlands (C)	7.94	10.29	7.52	17.80	8.45
Yorkshire & Humberside (C)	7.92	11.12	7.35	18.80	8.46
North West (B)	7.90	9.74	8.66	16.16	8.57
North (A)	7.88	9.11	10.25	21.36	8.64
Wales (W)	7.39	9.80	9.43	15.02	8.09
Scotland (x)	7.72	10.85	7.79	15.16	8.34

Source: authors' calculations from ARD.

Table 13: Percentage Distribution of Employment by Date and Enterprise Size

	Enterprise size				
Year of update	**0–9**	**10–19**	**20–99**	**100 or more**	**All enterprises**
1991	0.3	0.6	0.1	0.0	0.1
1992	0.0	0.1	0.0	0.0	0.0
1993	28.7	40.2	9.3	0.2	8.5
1994	0.2	0.4	0.1	0.0	0.1
1995	1.9	3.8	2.7	0.4	1.1
1996	1.6	4.7	4.7	1.0	1.8
1997	3.8	8.4	12.2	5.5	6.2
1998	3.0	8.2	32.6	12.7	13.2
1999	3.2	10.8	34.8	47.8	36.6
2000	0.4	1.1	3.4	32.4	21.7
Unproven enterprises	56.9	21.8	0.0	0.0	10.7
Total	100.0	100.0	100.0	100.0	100.0

Source: ARI, referring to the 2000 IDBR, cited in ONS (2001, Table 10, p.53)

Table 14: Basic data available for selected firms (year=2001)

	Non missing Observations	Mean	Median	Standard Deviation
Sales	6516	23358.51	285.28	373711.70
Gross Value Added	6446	5146.37	69.12	76672.22
Gross Output	6446	23687.73	294.35	376683.60
Dstocks	6516	60.99	0.00	2048.76
Capaq	6446	17.29	0.00	746.32
Ins	6446	7.89	0.00	208.15
Purch	6516	18353.33	202.00	300556.52
Emp	6024	315.41	5.00	4267.50
Part Timers	6520	172.19	2.00	2639.42

Source: authors' calculations from ARD.

Table 15 Productivity data

ARD data

Year	Gross output per person	Gross output per FTE	Value Added per person	Value Added per FTE	Frequency
1997	67.82	95.72	14.58	20.43	6,575
1998	66.87	94.07	14.25	19.79	6,436
1999	64.53	89.11	14.35	19.50	6,051
2000	73.65	104.51	15.69	22.14	6,110
2001	73.92	103.86	16.43	22.70	6,466
2002	75.22	106.76	16.45	23.06	6,167
Annual productivity growth, 1998-2002	3.12%	3.37%	3.86%	4.12%	
Annual productivity growth, 1998-2001	3.52%	3.47%	5.1%	4.89%	

Published

ABI, 98-2002 (retail)			3.87%		
Experimental (ret, distrib, hotel, catering)					
1997-98	0.20%	1.70%			
1998-2001	1.61%	2.75%			
EUKLEMS, 98-2001					
UK retailing			3.07%	2.31%	
US retailing			6.66%	7.17%	

Notes: see text for details. Upper panel are calculations based on ARD data. Lower panel are official published data. ABI is computed from the published ABI tables. Experimental are from Daffin, Reed and Vaze (2002). EUKLEMS are for US and UK from http://www.euklems.net/) which builds on the OECD STAN project.

Table 16: GVA Productivity, by employment size, 2002

All	**0-9**	**10-19**	**20-49**	**50-99**	**100-249**	**250+**
Emp	22.88	19.80	17.38	19.35	18.17	17.50
FTE	32.12	24.89	21.36	23.76	22.31	22.95
Frequency	3793	820	579	264	250	401

Non-specialised	**0-9**	**10-19**	**20-49**	**50-99**	**100-249**	**250+**
Emp	18.75	13.64	12.13	12.20	12.59	14.12
FTE	29.00	20.37	16.42	16.58	16.68	19.64
Frequency	622	98	74	42	41	103

Source: authors' calculations from ARD.

Table 17: Productivity spreads, 2001

Variable	Standard Deviation	I QR
Gva Per Head, unweighted	0.89	0.96
Gva per heard, weighted	1.23	1.26
Go per head, unweighted	0.75	0.9
Go per head, weighted	1.1	1.18
Frequency	5173	

Note: All data are transformed first into deviations from three digit industry means.
Source: authors' calculations from ARD.

Table 18: Results of the FHK decomposition, 1998-2002, by industry

	No Sampling Weights			With Sampling Weights			Frequencies			IndWeights
Industry	**Stayers**	**Entry-Exit**	**Productivity growth**	**Stayers**	**Entry-Exit**	**Productivity growth**	**Stayers**	**Exitors**	**Entrants**	
521. Non-specialised	-1.16	0.16	-0.05	2.14	-1.14	-0.02	104	287	203	0.566
522. Food, beverages, tobacco	2.28	-1.28	0.07	0.56	0.44	0.24	49	341	170	0.024
523. Pharmaceutical	0.33	0.67	0.34	0.28	0.72	0.36	27	94	73	0.023
524. Other retail	0.50	0.50	0.26	0.42	0.58	0.29	223	822	644	0.338
525. Second-hand	<0.40	>0.60	0.22	<0.10	>0.90	0.53	<5	55	<35	0.003
526. Not in stores	1.97	-0.97	0.12	3.85	-2.85	0.05	16	99	105	0.041
527. Repair	<0.30	>0.70	0.06	<-0.20	>1.20	-0.09	<10	28	<100	0.005
52. All Sectors	0.38	0.62	0.06	0.25	0.75	0.09	451	1726	1300	

Dropping top two RUs	**No Sampling Weights**			**With Sampling Weights**			**Frequencies**		
Industry	**Stayers**	**Entry-Exit**	**Productivity growth**	**Stayers**	**Entry-Exit**	**Productivity growth**	**Stayers**	**Exitors**	**Entrants**
521. Non-specialised	-2.08	-1.08	-0.01	-0.34	1.34	0.02	102	287	203
522. Food, beverages, tobacco	0.89	0.11	0.25	0.50	0.50	0.35	49	339	169
523. Pharmaceutical	0.62	0.38	0.41	0.43	0.57	0.42	27	92	71
524. Other retail	0.43	0.57	0.30	0.37	0.63	0.32	221	822	644
525. Second-hand	>2.00	<-1.00	0.14	<0.30	>0.70	0.50	<5	53	<35
526. Not in stores	0.44	0.56	0.05	0.57	0.43	0.02	15	98	104
527. Repair	<0.50	>0.50	0.08	<-0.20	>1.20	-0.14	<10	26	<100
52. All Sectors	0.42	0.58	0.10	0.30	0.70	0.13	449	1726	1300

Note: Productivity is calculated as gross value added per FTE. Numbers in STAYERS and ENTRY/EXIT columns are shares of total productivity growth in the "Prod g" column, except where "Prod G" is negative in which if the STAYER or ENTRY/EXIT effect is positive then the number in the cell is positive (to denote that the contribution to productivity growth is positive, even though productivity growth is overall negative). All sector data is industry share weighted STAYERS and ENTRY/EXIT effects divided by industry share weighted productivity growth, see text for details. Source: authors' calculations from ARD.