

# Accounting for growth in retail trade: an international productivity comparison

Robert Inklaar · Marcel P. Timmer

Published online: 12 September 2007  
© Springer Science+Business Media, LLC 2007

**Abstract** In this paper we argue that the standard approach for measuring output and productivity in the trade sector has become obsolete. The key problem is that changes in prices of goods purchased for resale are not accounted for. We outline a consistent accounting framework for measuring trade productivity and provide new estimates, taking into account purchase prices of goods sold in a double deflation procedure. We find strong productivity improvements in the UK and US compared to France, Germany and The Netherlands since The mid-1990s. This finding is robust for various productivity measurement models.

**Keywords** Productivity · Retail trade · International comparisons · Double deflation

**JEL Classifications** C82 · L81 · O47

## 1 Introduction

Most sectoral productivity studies have focused on firms or industries in the manufacturing sector. Recently, there is increased attention for service industries, in particular the distributive (wholesale and retail) trade sector. The main reason is that performance in the trade sector is one of the main drivers of the US growth acceleration in the 1990s (Jorgenson et al. 2005; Triplett and Bosworth 2004). Moreover, over half of the economy-wide labor productivity growth leads of the US over Europe after 1995 can be

traced to strong US performance in trade (van Ark et al. 2003). Various studies have discussed the reasons for superior performance in the US trade industry. Most focused on the role of ICT as a source of productivity growth and stress Europe's lagging behind in ICT investment and accompanying organizational changes. Others stress the restrictive European regulations concerning for example zoning and labor markets, and cultural differences (McGuckin et al. 2005; Gordon 2004).

In this paper we take one step back and raise the more down-to-earth question whether superior US productivity growth is real, or due to inadequate measurement. The latter possibility has been suggested by a recent report of the European Commission (2004) and was also raised by Gordon (2004). Both stress the statistical problems with measures of productivity in trade sectors and argue for a critical assessment of the way in which volume measures of trade output are being calculated. The European Commission report claims that the contribution of trade sectors to the US productivity boom is substantially overestimated. Despite some useful clues from recent studies on US measurement of retail service output (Triplett and Bosworth 2004; Manser 2005) this subject has not been investigated in-depth in an internationally comparative perspective. In this paper we argue that national accounts data are becoming more and more obsolete, due to organizational changes in retail, and suffer increasingly from comparability problems, due to statistical measurement innovations. The key problem is that in current National Accounts methodology, changes in prices of the most important input in retail trade, namely the purchases of goods for resale, are not accounted for. We outline a consistent accounting framework for measuring trade sector productivity and illustrate this with alternative measures of trade productivity in the US and four major European countries.

---

R. Inklaar · M. P. Timmer (✉)  
Groningen Growth and Development Centre,  
Faculty of Economics, University of Groningen,  
P.O. Box 800, 9700 AV Groningen, The Netherlands  
e-mail: m.p.timmer@rug.nl

The failure to account for changes in purchase prices is becoming more pressing for two reasons. First, changes in the business models of retailers are changing the demarcation between activities of traders, manufacturers and customers. Triplett and Bosworth (2004) provide a simple example regarding the sale of bicycles, which once were delivered to the retailer fully assembled. Now they typically arrive in a box, and customers can choose between having the store arrange for assembly and doing it themselves. Failure to account for differences in purchase prices can lead to misstated growth if certain activities are shifted between stores and suppliers (Triplett and Bosworth 2004; Manser 2005).

The second reason is that the use of quality-adjusted prices for deflation of goods sold and the lack of deflation of goods purchased, may lead to a bias in productivity growth in the trade sector. This is similar to the point stressed by Triplett (1996) in his study of productivity growth rates in computer manufacturing. Given that the use of quality-adjusted prices by statistical offices is rapidly increasing, this problem is becoming more important. Presently, it appears most visible in measurement of computer sales. For example, nominal sales of electronics stores grew on average at 5% per year in the US during 1995–2002. The prices of these products, about half of which are computers, declined on average at an annual rate of 12% as measured by quality-adjusted price indices. As a result, sales volume grew by a phenomenal 17% annually. In the national accounts, margin volumes are assumed to grow at the same 17% rate. However, the purchasers' prices of these computers will have declined substantially as well, so as Triplett (1996) argued, quality changes should be accounted for consistently throughout the productivity measurement framework. How this affects the comparability of productivity estimates across countries is still a black box, and has not gone much beyond speculations so far. In this paper we will provide a first attempt to quantify the magnitude of these measurement issues for international comparisons of productivity in retail trade industries.<sup>1</sup>

The remainder of this paper is organized as follows. In Sect. 2 we lay out a conceptual framework for measures of productivity in trade industries based on a neo-classical model of production and various concepts of output: sales, margins and value added. It is shown that multifactor productivity measures based on the various output concepts differ only by a scaling factor, which is proportional to the share of value added and margin in sales. However, this is

only true when all inputs are deflated with appropriate price indices. This is not the case in standard national accounting. In Sect. 3, we discuss the scope of the comparability problem for retail output and productivity measures between the US and Europe. In Sect. 4, the measurement problem is dealt with more directly by estimating double deflated margin volume measures. We provide an experimental approach to measure trade margins in constant prices, by estimating prices for both sales and purchases of goods sold. This is done for the US and four large European countries: France, Germany, The Netherlands and the United Kingdom. In Sect. 5 international comparisons of multifactor productivity are made, based on our internationally harmonized measures of trade output and inputs. Our main findings are first, that improved measures of retail trade output and productivity are feasible and second, that US productivity growth in the retail trade sector since the mid 1990s is well above continental European growth rates, and on par with growth rates in the UK. Section 6 concludes.

## 2 Models for multifactor productivity measurement

Productivity is, in general, defined as a ratio of outputs over inputs. Basically, one can choose between three output concepts in the trade sector: sales, margin and value added. This is different from other sectors (e.g. manufacturing) for which only a distinction is made between production and value added.<sup>2</sup> Margins are defined as the difference between the value of the goods sold (sales) and the value of the goods that would need to be purchased to replace them.<sup>3</sup> Gross value added is derived by subtracting costs of intermediate inputs from gross trade margins. Value added consists of compensation for labor and capital inputs. The relationship between the various output concepts is depicted in Fig. 1.

We start with setting up a general model of producer behavior in the trade sector following Jorgenson, Gollop and Fraumeni (1987). A full model of production would give the quantities of sales ( $q^S$ ) as a function  $F$  of quantities of goods purchased for resale ( $q^C$ ), intermediate inputs ( $q^H$ ), capital input ( $q^K$ ), labor input ( $q^L$ ) and technology, indexed by time  $T$ .

$$q^S = F(q^C, q^H, q^L, q^K, T) \quad (1)$$

Assuming a translog production function with constant returns to scale, the necessary conditions for producer

<sup>1</sup> The problem bears resemblance to the one that received a lot of attention in comparative studies of aggregate economic growth in the 1990s. Wyckoff (1995) pointed to the fact that deflators of ICT-investment varied across countries due to differences in the methods used to adjust prices for quality changes, and showed the consequences for labour productivity comparisons.

<sup>2</sup> For simplicity we omit the role of inventory changes. Diewert and Smith (1994) provide a firm-level model, which explicitly takes account of inventory. However, data on inventory changes are not available at the sectoral level for all the countries we analyze.

<sup>3</sup> According to the System of National Accounts (SNA 1993).

Sales of goods	Gross margin	Value added	Labour
		Intermediate inputs	Capital
Purchases of goods sold			

**Fig. 1** Output, margin, input and value added concepts in distributive trades

equilibrium lead to the following discrete approximation of the change in sales quantities:

$$\dot{q}^S = \dot{A}^S + w^L \dot{q}^L + w^K \dot{q}^K + w^H \dot{q}^H + w^C \dot{q}^C \tag{2}$$

This is the well-known growth accounting identity with  $\dot{q}$  denoting a (logarithmic) quantity change and  $w^x$  the share of each input  $X$  in total sales:  $w^L + w^K + w^H + w^C = 1$ , and  $A^S$  multifactor productivity (MFP) which is a measure of technological change ( $\delta F/\delta T$ ).<sup>4</sup> The weights are period averages. MFP growth rates can be derived as a residual by:

$$\dot{A}^S = \dot{q}^S - w^L \dot{q}^L - w^K \dot{q}^K - w^H \dot{q}^H - w^C \dot{q}^C \tag{3}$$

It shows that MFP growth is the difference between the sales quantity growth and a weighted average of the growth in quantities of primary factor inputs (capital and labor), quantities of intermediate inputs and quantities of goods purchased for resale. However, there is no data readily available to implement this model, as volume changes of goods purchased for resale are not recorded in official statistics. In this paper, estimates will be made for the first time (see Sect. 4). Instead, existing growth accounting studies such as Jorgenson et al. (2005) and Triplett and Bosworth (2004) take the margin as the trade output concept, where retail services are produced by using factor and intermediate inputs:

$$q^M = G(q^H, q^L, q^K, T) \tag{4}$$

This margin model is a restricted version of the full production model. Effectively one assumes separability

<sup>4</sup> In practice, the MFP measure also includes other elements besides technological change, such as efficiency changes, changes in input utilization rates, economies of scale etc. (see OECD 2001).

between goods purchased for resale and the various other inputs. It does not allow for substitution possibilities between capital, labor and intermediate inputs on the one hand, and goods purchased on the other. Under the same assumptions as above, a MFP measure based on margin output ( $A^M$ ) can be derived as follows

$$\dot{A}^M = \dot{q}^M - v^L \dot{q}^L - v^K \dot{q}^K - v^H \dot{q}^H \tag{5}$$

with  $v^x$  the share of each input in total margin:  $v^L + v^K + v^H = 1$ . In theory, MFP based on margins or sales differ only by a scalar as long as the margin is measured by means of double deflation. That is, when margin growth is measured as the difference between growth of sales and growth of goods purchased, each deflated by their own price index:

$$\dot{q}^M = \frac{1}{w^M} (\dot{q}^S - w^C \dot{q}^C) \tag{6}$$

with  $w^M$  the share of margin in total sales,  $1 - w^C$ . In this case there is a simple relationship between MFP measures of the full (sales-based) and the restricted (margin) model. Substituting (6) in (5) and using (3) one can easily show that they differ only by a scalar:<sup>5</sup>

$$\dot{A}^M = \frac{1}{w^M} \dot{A}^S \tag{7}$$

This only holds in case margins are double deflated. As stated before, neither quantities nor prices of goods purchased for resale are extensively collected within the official statistical systems. Instead, standard statistical practice is to derive real margin growth by assuming that the volume of margins follows the volume of sales:  $\dot{q}^S = \dot{q}^M$ .<sup>6</sup> In this case, the MFP measure based on national accounts data ( $\dot{A}_{NA}^M$ ) is as follows:

$$\dot{A}_{NA}^M = \dot{q}^S - v^L \dot{q}^L - v^K \dot{q}^K - v^H \dot{q}^H \tag{8}$$

The national accounts based MFP measure will be biased, depending on the difference in growth in sales and margin volumes. The difference between the two MFP measures follows from comparing Eqs. 5 and 8:

<sup>5</sup> This is analogue to the relationship between MFP measured on a gross output and a value added basis as outlined by Baily (1986).

<sup>6</sup> Although there were a wide variety of methods in use in OECD countries in the estimation of real trade output up to the beginning of the 1990s (see OECD 1995), a rapid convergence of methods has taken place since. Nowadays most countries in the OECD apply this methodology for measuring margin volumes. A minor part of trade output is represented by the output of specific services, e.g. repairing and intermediating services. These services are deflated directly by a corresponding price index. Oi (1992) and Triplett and Bosworth (2004, Chap. 8) provide a specific review of measurement issues in US retailing and Eurostat (2001) provides a useful general discussion of the problems in measuring real output in the trade sector.

$$\dot{A}_{NA}^M - \dot{A}^M = \dot{q}^S - \dot{q}^M = \frac{w^C}{w^M} (\dot{q}^C - \dot{q}^S) \tag{9}$$

When the quantity of goods purchased for resale ( $\dot{q}^C$ ) grows slower than the quantity of sales ( $\dot{q}^S$ ), MFP growth based on national accounts data is an underestimation. This could happen if certain activities are moved from the supplier to the store, such as in the case of the bicycle assembly example from the introduction. In addition, if quality improvements lead to a rapid fall in sales prices such as with computers, the quality improvements will lead to higher MFP growth under the current National Accounts methodology. However, since purchases prices will also be falling for those products, double-deflated margin growth will not suffer from this problem. In this paper we will study whether these biases exist for the retail sector and to what extent they differ across countries, by measuring not only growth in sales volumes, but also the growth in the volume of goods purchased for resale.

A third alternative model to measure multifactor productivity is based on value added. It is given by

$$q^{VA} = H(q^L, q^K, T) \tag{10}$$

Under the standard assumptions, MFP based on value added is given by

$$\dot{A}^{VA} = \dot{q}^{VA} - u^L \dot{q}^L - u^K \dot{q}^K \tag{11}$$

with  $u^L$  the share of labor compensation in value added, and similarly for capital. As in the case of the margin model, when the change in value added volume ( $q^{VA}$ ) is measured by means of double deflation according to

$$\dot{q}^{VA} = \frac{1}{v^{VA}} (\dot{q}^M - (1 - v^{VA}) \dot{q}^H) \tag{12}$$

with  $v^{VA}$  the share of value added in margin, there is a simple link between MFP growth in the value added model and MFP based on the other models as follows:

$$\dot{A}^{VA} = \frac{1}{v^{VA}} \dot{A}^M = \frac{1}{w^{VA}} \dot{A}^S \tag{13}$$

with  $w^{VA}$  the share of value added in sales.

Finally, output and inputs are aggregates from more detailed data. Growth in aggregate sales can be written as a weighted average of sales growth of detailed goods  $j$  ( $\dot{q}_j^S$ ) as follows

$$\dot{q}^S = \sum_j w_j^S \dot{q}_j^S \tag{14}$$

with  $w_j^S$  the share of product  $j$  in total sales. Similarly, growth in aggregate input X (capital, labor, intermediate inputs or goods purchased) can be written as a weighted average of detailed input  $x$  growth ( $\dot{q}_x^X$ ) as follows

$$\dot{q}^X = \sum_x w_x^X \dot{q}_x^X \tag{15}$$

with  $w_x^X$  the share of input  $x$  in total input X costs:  $\sum_x w_x^X = 1$ . Examples of detailed inputs include various types of capital (ICT and non-ICT asset types) and intermediate inputs (wrapping paper, advertising, legal services etc.). This will be used in Sect. 5 when calculating MFP.

### 3 The potential bias in productivity comparisons based on national accounts data

In the previous section we showed that MFP measures based on national accounts data could be biased due to the lack of data on prices and quantities of goods purchased for resale. This problem is especially pressing when the quality of goods is rapidly increasing and picked up by the sales price measures of the statistical office. In the US statistical system there has been a rapid increase in the use of hedonics in the quality-adjustment of price indices, especially, but not solely, for high-tech goods. This has led to dramatic price declines, for example in the case of computers. This poses problems for international comparisons of output and productivity. Measured sales volumes will be much smaller in countries that do not make use of hedonic methods for measuring IT-goods prices.

In Table 1 we provide an assessment of the potential impact of the use of hedonic deflators for IT-goods sales on the comparative volume measures of sales in the trade sector in the US and four EU countries (France, Germany, The Netherlands and the UK) during 1995–2002. Table 1 shows the average shares of ICT goods in total consumption of goods by households in these five countries. This share is a good proxy for the share of ICT goods in retail

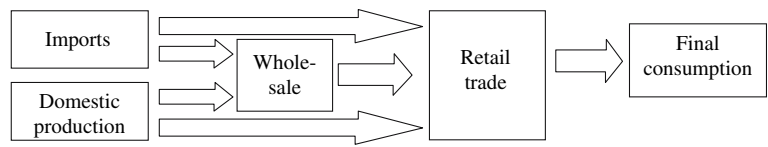
**Table 1** The impact of ICT goods sales on the growth of sales volumes in retail trade, 1995–2002

	ICT goods		Sales volume growth	
	Sales share	Sales price change	All sales	Excluding ICT goods
France	5.0	-5.8	2.6	2.2
Germany	3.0	-5.1	1.6	1.5
The Netherlands	3.4	-4.2	1.9	1.6
UK	7.0	-7.6	5.2	4.5
US	5.0	-14.0	4.5	3.7

*Note:* ICT goods cover the consumption of ‘Audio-visual, photographic and information processing equipment’ in total consumption of goods by households

*Sources:* National Accounts and other national sources

**Fig. 2** Simplified flow of goods between retail trade and production sectors



sales. The share of ICT goods consumption in the US is not particularly large, and in between that of the European countries. The second column, however, shows that recorded prices of ICT goods have dropped up to two to three times faster in the US than elsewhere. This confirms the suspicion raised by Gordon (2004) and the European Commission (2004) that the scope for an upward bias in the measurement of sales volume is bigger in the US than in Europe.

A straightforward way to obtain an impression of the size of the bias in trade sales due to the use of hedonic ICT goods deflators is by simply removing those retail and wholesale industries which mainly sell ICT goods and compare only the sales volume of non-ICT goods trading industries. This is done by for example Manser (2005) and the European Commission (2004). In columns 3 and 4 of Table 1 we provide sales volume measures of the retail sector, including and excluding sales of ICT goods. They show that ICT sales have been a major driver of retail sales in the US and more so than in other countries. For example, retail sales grew on average 2.9% faster in the US than in Germany, but only 2.2% when ICT sales are excluded. One should bear in mind that such exercises will provide an upper bound estimate of the bias, basically assuming that trade industries that sell ICT do not make an above-average contribution to sales growth, which is very unlikely given the rapid sales of ICT goods in the last decade.

In fact, international differences in deflation methods are not only confined to ICT goods. The problem is of a more general nature: when the measured change in sales volume is mainly due to changes in the quality characteristics of the traded goods rather than the physical volume of their sales, and countries differ in their quality adjustments, international comparability is compromised. This is true for ICT sales, but also true for sales of other goods whose prices are constant quality measures. Given the increased attention of statistical agencies for improved measurement of quality changes in price indices, this problem is getting worse over time. To the extent that the US is more advanced in terms of developing quality-adjusted price indices the upward bias in US trade sales volume measures compared to Europe may go beyond the bias created by recorded sales of ICT products. One way to assess the bias is by developing double deflated measures for output in trade industries. By combining various data sources, double deflation of retail margins is feasible. This is the topic of the remainder of this paper.

#### 4 Double-deflated measures of trade margins

For double deflation of retail margins, two sets of prices are needed: retail sales prices and retail purchase prices. The main problem is the derivation of retail purchase prices. Retailers purchase goods for resale mainly through wholesalers. But increasingly, the wholesale sector is bypassed and goods are acquired directly from domestic and foreign manufacturers. In Fig. 2 we provide a stylized view of the flow of goods through the retail trade sector.

We define a price for purchased goods by the retailer through matching a producer price index (PPI) and an import price index (IPI) to each final consumption good category  $i$ . For the retail trade sector, the change in the purchase price ( $\dot{p}^C$ ) is calculated as:

$$\dot{p}^C = \sum_i [w_i^I \dot{p}_i^I + (1 - w_i^I) \dot{p}_i^D] \tag{16}$$

with  $\dot{p}_i$  denoting a price change of product category  $i$ ,  $w_i^I$  the share of imports in total purchases and superscripts  $C$ ,  $I$  and  $D$  denoting respectively total purchases, imported purchases and domestic purchases. The share of imports in total purchases of each type of good is determined using input–output tables, under the assumption that the share of each goods category in total purchases equals the share of each category in total consumption.<sup>7</sup> Although the estimated sales and purchases prices correspond to household consumption on goods, one may assume that these are the relevant prices for the retail trade sector in each country. Note that by defining the retail purchase price as a weighted average of domestic production and import prices, we ignore the wholesale sector. To be more precise, we assume that changes in wholesale sales prices are proportional to changes in wholesale purchase prices (see below for more discussion). Prices of retail sales and purchases can be used to derive implicit prices for margins ( $\dot{p}^M$ ) using the dual price equivalent of Eq. 6:

$$\dot{p}^M = \frac{1}{w^M} (\dot{p}^S - (1 - w^M) \dot{p}^C) \tag{17}$$

with  $w^M$  the share of the margin in total sales.

<sup>7</sup> Note that with this assumption we make a link between products purchased for resale and products sold. Because we do not have purchase shares, we have to rely on sales, proxied by consumption, shares instead. Hence, the prices of products purchased are not independently measured from prices of sales, as they ideally should.



**Table 2** Sales, purchases and margin prices for household consumption goods

	Sales prices	Purchase prices	Purchase price contribution by		Implicit Margin prices
			Domestic products	Imports	
<i>1987–1995</i>					
France	1.9	0.7	0.9	–0.2	4.3
Germany	1.7	1.3	1.3	–0.1	2.5
The Netherlands	1.0	0.4	0.5	–0.1	2.1
UK	3.3	4.1	3.2	0.9	1.6
US	2.0	2.2	1.7	0.5	1.7
<i>1995–2002</i>					
France	1.4	0.5	0.7	–0.2	3.4
Germany	0.7	1.0	0.9	0.1	0.3
The Netherlands	1.5	1.1	1.0	0.0	2.2
UK	0.2	0.3	0.7	–0.5	0.3
US	0.5	0.7	0.9	–0.2	0.0

*Note:* Household consumption on goods excludes automobiles and gasoline consumption

*Sources:* Sales prices: household consumption expenditure prices (National Accounts); Import prices of purchases: import price index (National Accounts and other national sources); Domestic prices of purchases: gross output and producer price indexes (National Accounts and other national sources); Import shares in purchases: Input-Output tables from national sources, Retail margin-to-sales ratios: Timmer and Ypma (2006) for 1997 extrapolated on basis of retail census and surveys

Our empirical analysis covers France, Germany, The Netherlands, United Kingdom and the United States. For each of these countries, National Accounts statistics provide information on household consumption by type of goods in current and constant prices. The level of detail varies between countries, but we could use consumption data for in between 20 and 40 goods categories, such as food products or clothing. Our measure of the margin to sales ratio is based on a benchmark estimate of retail margin to sales ratios derived from industry surveys and censuses for the retail sector for each of the countries for 1997 (Timmer and Ypma 2006). These estimates are extrapolated using gross margin to sales ratios from census sources.<sup>8</sup>

In Table 2 we present the results of our double deflation procedure for retail margins.<sup>9</sup> The table presents the sales price of household consumption goods, the corresponding purchase prices and the contributions from import and domestic price developments for the period from 1987 to 1995, and the period from 1995 to 2002. We also present the double-deflated margin prices using the retail trade margin-to-sales ratios as weights according to Eq. 17. The table reveals considerable heterogeneity between the countries as well as between the two periods. In the period

1987–1995, retail sales prices in the UK and the US grew faster than in the rest of Europe. This was driven by much more rapid growth in purchase prices of both domestically produced goods and imports. But after 1995, this was reversed. Anglo-Saxon sales prices grew less than elsewhere, thanks to a sharp decline in purchase prices (especially imports). When looking at the development of the margin prices, a distinction can be made between Germany, the UK and the US on the one hand, and France and the Netherlands on the other hand. Whereas in the first group of countries, the margin prices decelerated, this was not, or much less so, the case in the other countries.

Using double-deflated margin prices, estimates of the growth rates of margin volumes in retailing can be computed according to Eq. 6. These are shown in Table 3, together with sales volume growth rates for 1987–1995 and 1995–2002. In the case of Germany, The Netherlands, the

**Table 3** Average annual growth of real retail sales and margins, 1987–2002

	1987–1995		1995–2002	
	Sales	Margins	Sales	Margins
France	2.4	0.1	2.6	–0.2
Germany	2.9	1.4	1.6	2.4
The Netherlands	2.3	2.6	1.9	2.2
UK	3.0	6.3	5.2	6.5
US	2.8	3.1	4.5	4.9

*Sources:* see Table 2

*Notes:* Germany refers to 1991–1995 instead of 1987–1995

<sup>8</sup> If no reliable data on retail sales could be found for some years, sales were extrapolated using the growth rate of household consumption on goods.

<sup>9</sup> Note that we use the ISIC rev 3 definition of retail trade (ISIC 52). This industry does not include retail trade of automobiles and gasoline. Hence these items are also excluded from household consumption.

**Table 4** Sources of growth in retailing, 1995–2002

	France	Germany	Netherlands	UK	US
Growth of double deflated margin	-0.2	2.4	2.2	6.5	4.9
Contribution of growth					
Intermediate inputs	0.7	0.5	1.1	2.1	1.0
Capital services	0.6	0.3	0.3	0.6	0.7
Hours worked	0.2	-0.1	0.6	0.5	0.5
MFP	-1.6	1.7	0.2	3.4	2.8

*Notes and sources:* Multifactor productivity measures based on the margin model, see Eq. 5; Based on double-deflated margins from Table 4 and intermediate input growth from GGDC (2004) for France, Germany and The Netherlands, and for UK and US as described in the Appendix. Price change of intermediate inputs in Germany is equal to the price change of intermediate inputs for total trade; Hours worked from GGDC (2004) 60-industry database. US labor input is adjusted by adding an estimate of hours worked in headquarters for trade industries; Capital service growth from Inklaar et al. (2005)

UK and US, margin volumes grow faster than sales volumes, albeit the difference is relatively small in the most recent period. This is mainly due to increasing nominal margin-to-sales ratio and slow growth of margin prices. For example, German margins grew almost 1%-point faster than sales. Table 2 shows that increases in prices of goods purchased by German retailers (up by 1.0%) have not been fully passed through to the final consumer: sales price went up 0.7%. This is indicative of improved efficiency in German retailing, or a long-term squeeze of retailing profits. For France, the opposite is found: margin volumes grew much slower than sales. In contrast to Germany, sales prices in France have grown much faster than purchase prices, which is indicative of a rapidly increasing price for retail services. Our double deflated margin measure corrects for this service price increase, whereas the sales measure does not. Summarizing, this exercise would suggest that output comparisons based on sales data do not upwardly bias US growth when comparing it to the European countries, except for the UK.

The usefulness of double deflation depends critically on the availability and quality of price indices for goods purchased. In particular, these prices need to be adjusted for quality in the same way as prices of goods sold. Normally, statistical offices put more effort in quality adjusting consumer prices, which is the basis of our output price measure, than in producer and import prices, which are the basis of our goods purchased prices. This will be true for all countries and more detailed consumer, producer and import prices might alleviate this problem. But no systematic bias is to be expected.<sup>10</sup> Another potential measurement error in our procedures is that we allocate all the change in margin prices to retailing, ignoring the role of changes in wholesale margins. On average, wholesale margins make up about a quarter of the total margin on consumer goods, so

the potential effect is limited. But without data on wholesale sales prices, this issue cannot be resolved.

## 5 Multifactor productivity comparisons

In this section we provide a breakdown of the sources of growth in retailing for our set of countries and calculate various MFP measures. In order to do this, we had to solve a number of other comparability problems first. These had to do with the measurement of intermediate inputs in the national accounts of the UK and the US. In the UK, no estimates are made of volume changes in intermediate inputs (Sharp 2003). We prepare new estimates of intermediate input use in UK trade and find that the official value added growth rates for retail are only slightly over-estimated (4.6% vs. 4.4%). For the US we made adjustments to bridge the European industrial classification NACE and the North American industrial classification (NAICS). This involved the reallocation of trade industries and of value added for the inputs from the sector ‘management of companies’.<sup>11</sup>

In Table 4, growth of our double deflated margin is broken down into the contributions of growth from intermediate inputs, capital, labor and MFP, based on Eq. 5. Labor input is measured as hours worked and capital service growth rates are from Inklaar et al. (2005).<sup>12</sup> The latter study presents capital service growth rates based on 6 asset types, including ICT hardware and software for the trade sectors using Eq. 15. The table shows the contribution of each source of growth, which is calculated as the share of the factor in the margin times the growth of the factor. Capital service growth rates vary little across countries and

<sup>11</sup> See Timmer and Inklaar (2005) for details.

<sup>12</sup> In this paper we do not deal with issues concerning factor input measurement. Although there are problems in the international comparability of factor input measures in distributive trade too, they are relatively minor compared to problems with margin measures.

<sup>10</sup> See Timmer et al. (2005) for detailed studies of 12 US retailing industries.

are only a minor contributor to growth differences. The same is true for working hours, which even contributed negatively in Germany. Instead, MFP is the major driver of margin growth rates in Germany, the UK and the US, followed by growth in intermediate inputs. MFP explains more than half of the retail margin growth in these countries, but was negligible in The Netherlands and even negative in France in the period 1995–2002.

Finally, we look at the differences in comparative performance when alternative multifactor productivity measurement methods are used. As discussed in Sect. 2, multifactor productivity growth rates can be estimated on the basis of a full production model using sales as the output measure, or more restricted models based on margin or value added. In Table 4, we used the margin model. When margins and value added are measured on the basis of a double deflation procedure, MFP growth rates based on the various models are scaled estimates of each other. This is indicated in Eqs. 9 and 13. The ratio of MFP growth based on sales and MFP growth based on margins is the margin-to-sales ratio. Similarly, the ratio of margin-based MFP and value added-based MFP is given by the value added-to-margin ratio.

In the first three columns of Table 5 MFP rates based on the three alternative production models are presented. As to be expected, MFP growth rates based on sales are lower than those based on margins, which in turn are lower than those based on value added. The main conclusion to be drawn from these results is that the use of a particular production model is inconsequential for international comparisons of productivity. Margin-to-sales and value added-to-margin ratios do not differ greatly across countries, so the ranking of each country in terms of productivity performance is not affected: the US and UK are leading according to all models, while The Netherlands and especially France are clearly lagging behind.

## 6 Concluding remarks

In this paper we have argued that at times of rapid improvements in the quality of products and organizational

change within the retail sector, sales as a measure of trade services output is becoming increasingly obsolete. From a multifactor productivity perspective, deflation of sales by a quality-adjusted price index poses no particular problems as long as purchased goods are also measured in constant-quality terms. But in productivity studies, measures of goods purchased have not been taken into account so far. This study is the first to present estimates of retail trade productivity taking into account purchase prices of goods sold. We find strong multifactor productivity growth in the UK and US compared to France, Germany and the Netherlands since the mid-1990s. This finding is robust to various models of MFP measurement, based either on sales, margin or value added.

This line of research could be extended to a study of the wholesale sector. Most likely, productivity comparisons of the wholesale sector suffer from greater biases than of the retail sector. Wholesaling of ICT goods accounts for a larger share of total sales than retailing of ICT goods because wholesaling includes ICT-goods exports and deliveries of ICT to the business sector that invests in ICT. Both exports and deliveries to business are sizeable flows in addition to domestic retailing. Therefore it seems safe to say that in wholesaling productivity measures, which ignore purchased goods are much higher than margin based measures. Coupled with the fact that the ICT goods producing sector in the US is much bigger than in Europe, comparative performance in wholesaling will be much more sensitive to the exclusion of goods purchased than retailing comparisons.

Finally, it should be stressed that the estimates in this paper are of an experimental nature and mainly developed for assessing the feasibility of a more complete accounting system for retail trade. There are still important data issues to be resolved before such measures are part of the statistical routine, but our results do suggest that strong US productivity growth in retail trade is not just due to different measurement practices. It should be pointed out that neither current national accounts methods nor our data exercise are able to deal directly with improvements in service *quality*. The quality of services is dependent on

**Table 5** Alternative estimates of multifactor productivity growth, 1995–2002

	Multi factor productivity growth based on			Average share of value added in	
	Sales	Margin	Value added	Margin (%)	Sales (%)
France	-0.5	-1.6	-2.0	68	16
Germany	0.6	1.7	2.7	62	14
The Netherlands	0.1	0.2	0.3	60	13
UK	1.1	3.4	5.8	51	13
US	0.9	2.8	4.2	65	16

Sources: see Table 5, based on Eqs. 3, 5 and 11



store characteristics like the convenience of the location of the store, the variety of goods on offer, information and swiftness of service, but also includes ancillary services such as credit facilities, delivery, after-sales service etc. Econometric studies have tried to measure this bias and generally conclude that trade service quality has improved in the US (see e.g. Betancourt and Gautschi 1993). However, comparable studies on these issues for Europe have not been made so far. Currently some experimentation is going on with collection of product margin prices by asking stores the difference between the sales and purchase price of a particular product, but on a very limited scale.<sup>13</sup> One way to measure quality change would be to directly measure price margins of specific items and to correct these for quality changes by collecting characteristics on the store in which the items are sold. As yet, current statistical practice is not well suited for studies of productivity growth in the distributive trade sector. The development of quality-adjusted price indices for margins or for goods purchased for resale would be a major step forward in this process.

**Acknowledgements** We are grateful to the BLS and BEA for helpful advice on the US data. We thank Bart van Ark for help and advice at various stages in developing this paper. We also received helpful comments from participants at a EU workshop on service statistics on 21 September 2004, at the final conference of the EPKE (Employment Prospects in the Knowledge Economy) project in London, 28–29 October 2004 and the OECD Workshop on Services, 15–16 November 2004, especially from Barry Bosworth, Marilyn Manser, Robert McGuckin and Werner Roeger.

## References

- van Ark B, Inklaar R, McGuckin RH (2003) ICT and productivity in Europe and the United States. Where do the differences come from? *CESifo Econo Stud* 49(3):295–318
- Baily M (1986) Productivity growth and materials use in US manufacturing. *Quart J Econ* 111:185–195
- Betancourt R, Gautschi D (1993) The outputs of retail activities: concepts, measurement and evidence from US census data. *Rev Econ Stat* 75:294–301
- Diewert WE, Smith AM (1994) Productivity measurement for a distribution firm. *J Product Anal*, 5:335–347
- European Commission (2004). The EU Economy 2004 Review, European Economy. No 6. 2004. Office for Official Publications of the EC. Luxembourg
- Eurostat (2001), Handbook on price and volume measures in national accounts, European commission, Luxembourg
- Groningen Growth and Development Centre (2004), *60-industry database*, October 2004 at www.ggdc.net
- Gordon RJ (2004) Why was Europe left at the station when America's productivity locomotive departed? CEPR Discussion Papers, no. 4416, London
- Inklaar R, O'Mahony M, Timmer MP (2005) ICT and Europe's productivity performance industry-level growth account comparisons with the United States. *Rev Income Wealth*, 51(4):505–536
- Jorgenson DW, Gollop FM, Fraumeni BM (1987) Productivity and US economic growth. Cambridge MA: Harvard University Press
- Jorgenson DW, Ho MS, Stiroh KJ (2005) Information Technology and the American Growth Resurgence, MIT Press
- Manser M (2005) Productivity measures for retail trade: data and issues. *Montly Labor Rev* 30–38
- McGuckin RH, Spiegelman M, van Ark B (2005) Perspectives on a global economy: the US advantage in trade performance. How can Europe catch-up? The Conference Board
- OECD (1995) Services: measuring real value added, OECD, Paris
- OECD (2001) OECD productivity manual: a guide to the measurement of industry-level and aggregate productivity growth, OECD, Paris
- Oi W (1992) Productivity in the distributive trades: the shopper and the economies of massed reserves. In: Griliches Z (ed) *Output Measurement in the Service Sectors*. University of Chicago Press
- Sharp P (2003) Gross domestic product: output approach (gross value added)-Revised. National Statistics Methodology Series NSM 32, London: ONS
- Timmer MP, Inklaar R, van Ark B (2005) Alternative output measurement for the US retail trade sector. *Montly Labor Rev* 2–8
- Timmer MP, Inklaar R (2005) Productivity differentials in the US and EU distributive trade sector: statistical myth or reality? GGDC Research Memorandum GD-76, Groningen Growth and Development Center
- Timmer MP, Ypma G (2006) Productivity in distributive trades: new comparisons for OECD countries. GGDC research memorandum GD-83, Groningen growth and development center
- Triplet J (1996) High tech industry productivity and hedonic price indices. In: OECD, industry productivity. International comparison and measurement issues, OECD Proceedings, Paris, pp 119–142
- Triplet J, Bosworth B (2004) Productivity in the US services sector. New sources of economic growth. The Brookings Institution, Washington DC
- Wyckoff AW (1995) The impact of computer prices on international comparisons of labor productivity. *Econ Innovat New Technol* 3:277–293

<sup>13</sup> Recently, the BLS in the US has introduced a new initiative to measure margin prices in its PPI program by surveying directly the difference between the sales price of a specific item and its acquisition cost (Manser 2005). Also in Europe experimentation is taking place, for example in Finland and Norway (Eurostat 2001).