CLOSING THE GAP, CROSSING THE LEVELS

IDEAS FACTORY

MULTI-DISCIPLINARY LITERATURE REVIEW
## CONTENTS

### The Role of Management Practices in Closing the Productivity Gap  3 - 55

Uwe Aickelin, Giuliana Battisti, Helen Celia, Chris Clegg, Xiaolan Fu, Rafael De Hoyos, Alfonsina Iona, Adriano Peixoto, Peer-Olaf Siebers

### The Unintended and Indirect Effects of Performance Measurement and Regulation on Productivity  56 - 118

Joseph Antony, Gerben Bakker, Vlado Balaz, Naomi Brookes, Stavroula Ionopoulou, Rosalind Rae, Anjulo Gurtoo, Kim Tan, Kathryn Walsh, Allan Williams

### Moving to the Next Level: A Scoping Study  119 - 132

Guiliana Battisti, Neil Burns, Delores Anon Higon, Jillian MacBryde, Beverley Randle, Marek Szwejczewski, Elizabeth Watson

### Intra-Firm and Inter-Firm Knowledge Transfers and Productivity in the Retailing Sector  133 - 161

Jeremy Clegg, Odul Bozkurt, Irena Grugulis, Dolores Anon Higon, Nicholas Vasilakos, Allan Williams

### Understanding and Assessing the Concept of Knowledge Leakage  162 - 183

Paul Chan, Rhodes Coles, Lyn Dodds, Andrew Grantham, Raphael Kaplinsky, Souad Mohamed, Diane Mynors, Kathryn Walsh
The Role of Management Practices in Closing the Productivity Gap

Uwe Aickelin, Guiliana Battisti, Helen Celia, Chris Clegg, Xiaolan Fu, Rafael De Hoyos, Alfonsina Iona, Adriano Peixoto and Peer-Olaf Siebers

Abstract
Management practices are undoubtedly inextricably linked to a company’s productivity, however research findings are mixed. This paper provides a multi-disciplinary review of the current evidence of such a relationship and offers suggestions for further exploration. We first summarise recent work investigating the existence and size of the productivity gap between the UK and other countries. We then lead on to discuss how productivity and management practices have been measured. Research findings from studies trying to understand how Information and Communication Technology (ICT), operations, and Human Resource (HR) management practices (amongst others) impact upon productivity are discussed. Finally, we suggest the application of a new operational research methodology (agent based modelling and simulation) to provide new insights into the effect of management practices. The paper ends with a summary of findings.

1. Assessing the Gap
The productivity gap is identified and measured in many papers and reports using a range of distinct indices. At first glance and without careful scrutiny, many findings are ambiguous and even contradictory. A great deal of these apparent discrepancies are accounted for by the different metrics and time spans used, the sector that has been focused on and methodological differences in national account procedures. On the other hand, “there are a number of reasons why measured productivity may differ, which do not necessarily reflect underlying differences in productivity” (Griffith and Harmgart, 2005). Traditionally, international comparisons have been made in terms of labour productivity, which has the advantage of being an easy and simple measure to achieve but lately, a more widespread comparison under Total Factor Productivity (TFP) has been used (Crafts and O’Mahony, 2001).

In a study produced for the Department of Trade and Industry (DTI) about the state of UK competitiveness, Porter and Ketels (2003) identify, based on OECD figures, a gap in labour productivity of about 15 % with the United States, 11% with France and 8% with Germany. In the same year, in a study for the Advanced Institute of Management Research (AIM), Griffith et al. (2003) place the gap with the US at “just over 40%”, a difference of almost triple the size when compared with the previous study. This is a typical example of the problems involved in the gap estimation. Nevertheless, two important lessons can be learned from this disagreement: Firstly, they address the economy in different ways. While the former study refers to the total economy, the latter is related to the business sector only, in other words, it does not take into account public administration, health, education and property (ESRC, 2004). Secondly, the studies deal with diverse measures of the same phenomenon. While in the DTI paper the productivity gap is expressed in terms of output per hour worked in the AIM report the gap is calculated in terms of value added per worker.
In a report published by the HM Treasury (2000), the gap is estimated at about 19% with the US, 20% with France and 18% with Germany. Here once again, the way in which the gap is measured accounts for the difference in the size when compared to the studies previously quoted. In this case, the gap is a measure of TFP. However, the Treasury also deals with productivity gap estimations on Gross Domestic Product (GDP) per worker and GDP per hour worked. In these cases, there is a gap variation from somewhere slightly over 22% with US (GDP per hour worked) up to about 44% (GDP per worker). The same applies when we compare the figures of UK with either France or Germany but within much smaller amplitude on the gap size.

From the previous examples some partial conclusions can be drawn: multiple measures capture different aspects of this complex phenomenon, but at the same time this is also responsible for a good amount of noise in the area. There is no correct way of dealing with it, just a number of choices determined by researcher and policy maker objectives and beliefs. Hence, there is a range of findings about the existence of a productivity gap with the US, France and Germany, in particular the estimation of its size. For instance, the studies of O’Mahony and de Boer (2002) and O’Mahony and van Ark (2005) give support to a persistent gap in the UK’s labour productivity performance compared with the US, France and Germany. At the same time, according to Porter and Ketels (2003), over the period 1995-2001 the UK registers productivity levels comparable to Sweden, Italy, Finland and Spain.

However, an improved outlook for the UK’s productivity emerges when we look at the latest OECD data published in 2005 (OECD in Figures, 2005). The situation previously pictured does not hold. Instead of lagging largely behind Germany, the UK appears in a surplus position and the gap with France and the US is much narrower than suggested by the papers and reports previously mentioned. A possible explanation for these differences might be found in an association of economic and productivity growth over the last few years.

Yet another different picture emerges when instead of comparing GDP per worker the comparison is performed under GDP per hour worked. In this case, the UK lags the US while standing behind Germany and France, and France even has a higher productivity level when compared with the US (OECD estimates, 2005). However, another ranking is achieved when you look into the living standards, following the suggestions from O’Mahony, Oulton and Vass (1998), measured as GDP per capita. Here, the UK lags behind the US while standing in front of France and Germany measured in the amount of money available per person in that country. The general conclusion that we can reach at this point is that whilst working longer hours than France and Germany, British people have a higher level of disposable income.

Another important factor to look at in this context is the level of unemployment. There are some suggestions that when unemployed workers go into the market, they could increase country productivity, but this is very unlikely to happen, because one of the main characteristics of this contingent of the workforce is their relative lack of skills when compared to other workers already employed. Thus, their contribution to increasing the general level of productivity would be marginal, if any (OECD, 2005). Over the last decade, all countries, except for Germany, managed to decrease their unemployment rates (OECD in Figures, 2005), but the UK managed to do this to a much higher extend than the other countries. However, as the additional workforce represent mainly the less skilled workers, this is unlikely to increase the general productivity level of the UK.
Summarising this section the following conclusions can be drawn. Firstly, there seems to be enough empirical evidence to support the idea of a productivity gap but the size of it is a disputable question. As the gap can increase or decrease depending on the kind of metrics applied, a more or less dramatic picture of UK performance can be drawn. It is not advisable to underestimate the bias caused by methodological choices when evaluating the gap or a country’s performance. Secondly, the data shows a picture of a strong UK economy performing well over the last decade with higher productivity growth rates than compared to Germany and France and there is a visible catching up movement with the US. Thirdly, because of high employment rates, economic growth in the UK is more likely to happen because of improvements in productivity rates. This differs from Germany and France, which still have room to experience a cycle of economic growth through decrease of unemployment. Fourthly, different productivity levels do not necessarily express economy inefficiency or incapacities, but may be an expression of a certain economic environment characteristics subject to political trade off. It is possible to conclude that productivity is a more vital subject to UK than to France and Germany.

1.2 Setting the gap into context: a historical perspective

Another important aspect when referring to the productivity gap is brought to our attention in a paper from Broadberry and O’Mahony (2004). The authors stress the need to put differences with US and other European countries within a historical perspective. Comparing US/UK labour productivity levels by sector at an aggregate economy level for a period ranging from 1869 to 1990, they show that the US overtook the UK around the turn of the twentieth century and when compared to Germany the underperformance of the UK economy dates back to the late sixties. This analysis is echoed in work by Maddison (1982) quoted by Rupert (1995).

In the first instance, the productivity acceleration of the American economy might be explained by greater availability of land and natural resources coupled with scarcity of skilled workers leading to an intensive use of technology when compared with the UK, i.e. a unique combination of conditions available in the New World (Rupert, 1995). This persisted until the Second World War when European economies suffered a set back.

The post war period is characterised by two distinct approaches to economical organization: while France and Germany adopted a development model based on their needs and characteristics, the UK tried to follow an Americanised pattern without sufficient adaptation to local circumstances (Broadberry and O’Mahony, 2004). Although this hypothesis has not been thoroughly scrutinised, it could be an explanation for the differences in productivity growth rates for UK, France and Germany after the war. From the late sixties onwards, those choices had the opportunity to be put to the test when Japan emerged as a strong player in the economic global arena. At that time, French and German companies were better than the UK’s, because the way in which the French / German economy was organised was more suitable to adapt to new competition. Such considerations are driven, according to the authors, by the need to avoid economic fashions that from time to time appear on the economic and political agenda.

However, over the last decade or so, the economic and political environment has experienced another radical transformation on the way people, companies and government relate to each other. Some authors, like Sennet (2006), refer to this as the period of ‘New Capitalism’; others, for example Temple (2002) or Van Reenen (2001) call it a ‘New Economy’. In
They share the view of a world where production and organizations are built over flexible bases and the intangible assets are placed high in the value chain. In other words, a very different economic environment when compared to the one following the war. We could say that the original situation responsible for the superior French and German productivity performance after the war is not in place anymore. If this hypothesis is correct it should be possible over the last decade or so to spot stronger performance indices growth rates for the UK when compared to France and Germany and possibly with the US as well. However, the differences in size, structure and resources between France, Germany and UK when compared with the US are so big that sometimes it is worth asking if a comparison is actually sensible. An alternative approach would be to compare the US with the EU.

1.3 Measurement Aspects

It is known that when explaining a country’s productivity, and hence the productivity differentials across countries, an essential step of the analysis process consists of defining and accurately measuring productivity. Statistical agencies and academics typically use labour productivity, measured in terms of both output per worker or per hour worked, according to the idea that in the long run, growth depends on increases in employment rather than on intensity of work (HM Treasury, 2005; Reynolds et al., 2005, Porter and Ketels, 2003). However, a strand of research has shown a number of concerns about the measurement of productivity.

From a macroeconomic perspective, some studies have shown that the productivity estimations may be biased if outputs or inputs are not accurately measured under waves of Information Technology (IT) investments. For example, Basu et al. (2003) shows that when new technologies are introduced, a need of organizational change arises that may be modelled as the accumulation of intangible complementary capital. This means that the firm is producing a stream of intangible outputs that constitute gross investment in complementary capital. The problem is that some of this output is not typically measured in the national accounts, and at the same time investment in IT may be associated with lower TFP as resources are diverted to organizational changes and learning. Violante (2003), analysing the missing productivity growth in the UK (or the exceedingly high productivity growth in the US), has taken into account measurement problems in the TFP growth. The idea is that periods of strong investment in IT are times where mostly output is unmeasured, so the true TFP growth may be underestimated, whereas periods where the economy has large stocks of IT and complementary capital are times where inputs may be grossly under-measured and the true TFP growth may be overestimated.

From a microeconomic perspective, output is often measured by sales or value added per worker, or per hour worked, deflated by a price index. In other words, researchers and statisticians sample a large number of items and see how prices change over time. This raises the issue of price comparability across markets: whether prices are reliable indicators of the market value of goods, and whether all the goods may be priced. Recent studies by Griffith et al. (2003) and Griffith and Harmgart (2005) argue that, in competitive markets, prices are usually overstatement, i.e. they reflect not only the value of output or input but also the market power. If this happens in the output market, they lead to an overstatement of productivity and if it happens in the inputs market, they lead to an understatement of productivity. In this perspective, comparisons of productivity across sectors or countries may be biased.
Secondly, measuring and comparing productivity may be problematic because in some sectors goods have no prices. That is, although sampling a large number of items and seeing how prices change over time may be relatively easy to do for manufacturing, it seems that problems arise in defining the value of output in the services sector as in this sector many items - such as health - have no prices (ESRC Seminar Series, 2004; Reynolds et al., 2005). Another problem arises from the fact that, since those items are measured in inputs, comparing productivity across sectors may be problematic if the measurement of labour inputs is inadequate or different between sectors (Reynolds et al., 2005, Pilat, 2005).

1.4 Sector contribution to the productivity gap

Although, displaying strong productivity growth in previous years, this has not been enough for the UK economy to overcome the differences with the US. Current analysis shows that the largest contribution to the gap is from the services sector, updating long historical trends regardless considerable variation over the years on sectoral contributions to the gap.

In a early study, O’Mahony, Oulton and Vass (1998) found that the UK productivity was “lower than in the three other countries” with the gap, relative to Germany and France being higher in services than manufacturing but when compared to US, the UK performed better in services than in manufacturing. The apparent contradiction with Broadberry (1998), previously quoted, could be explained by the differences in time span between the two studies.

Over previous years, the UK has improved its performance and caught up with the US in manufacturing and nowadays a considerable amount of the gap refers to the service sector (Griffith et al., 2003; ESRC, 2004; Crafts and O’Mahony, 2001). Three areas stand out as the biggest individual contributors to the gap with US: wholesale and distribution, financial intermediation and machinery and equipment. Together, they are responsible for more than 50% of the total gap measured in value added per worker (Griffith et al., 2003).

Recent empirical data supports these observations. Data in the OECD Fact Book 2006 (2006) reveals that the US experienced a huge growth in its industry productivity rate especially during the period of 1995 to 2000. The difference when compared to the UK is bigger than 20%, showing that a gap was built over the last decade. The data is compatible with studies showing a huge acceleration of US productivity during the second part of the nineties. For instance, Oulton (2002) observes that “from 1995 onwards a striking and unexpected increase occurred in the growth rate of labour productivity...in the United States”. It also supports O’Mahony, Oulton and Vass (1998) when they assert a worse situation for the UK’s manufacturing sector when compared to US manufacturing. At the same time, services present a better situation (OECD Fact Book 2006). During the period when the US was experiencing a huge growth of its industrial productivity rates when compared to France, Germany and UK, the same could not be observed in services. In this case, the UK’s productivity growth was slightly stronger than that from the US and from Germany and was 27% points ahead of France.

How can these figures be interpreted? A first conclusion points towards strengthening previous data showing an improved situation when the UK is compared with France and Germany. The sector breakdown shows that the catching up movement with Germany happens because of improvements in productivity in both the industry and the service sectors.
With France, an improvement was achieved in services, but part of it was lost in manufacturing. Most of the differences with the US (taking only into account the business sector of the economy) over the period under consideration might be explained by differences in labour productivity in manufacturing.

The significance of these numbers is highlighted when we observe the weight services sectors have on the economy either in terms of contribution to the GDP or the number of workers employed. Figure 1 below shows the contribution of each sector to the total economy in the UK, the US, France and Germany. We can see that the UK’s service economy is bigger than Germany’s and is catching up with France and US in terms of importance supporting the analysis previously delivered.

**Figure 1. Sector contributions to gross value added. Source: National accounts of OECD countries. OECD, Paris, 2005**

Overall, the conclusions we can draw from studying the productivity gap literature are that there seems to be a gap between the UK and other countries, but the size, sectors and reasons for the gap are a matter of ongoing debate. As far as the service sector is concerned, the UK seems to perform as well as its competitors over the last few years. However, this is assuming that all countries started from the same level in 1990, whereas a gap may have already existed. Moreover, so far we have only talked about the service sector as a whole. We are particularly interested in the retail segment of the service sector and hence need to look more closely at the way productivity is measured there in the following sections.

### 2. Measurement Issues

#### 2.1 Productivity Measurement

**2.1.1 General measurement of productivity**

The economic literature has used two different definitions of productivity. The first and more straightforward one, called partial factor productivity, captures the amount of output per labourer or hours of work. Define $Y_i$ and $L_i$ as total output and labourers (or hours worked) in
firm $i$ respectively, labour productivity can be computed as a simple ratio of production per labour inputs:

$$y_i = \frac{Y_i}{L_i}$$  \hspace{1cm} (1)

The problem with this type of measure is that, as pointed out by economic theory, the marginal productivity of labour (or any other factor of production) changes with the amount of other substitutes/complementary inputs available in the firm. For example, given that capital and labour enter as complement inputs in the production function, an extra machine will increase labour productivity, everything else constant. Moreover, the labour productivity profile is a decreasing function of the number of labourers, in other words, increasing the number of workers would decrease their average productivity, ceteris paribus. Therefore labour productivity gives only partial information of how inputs are transformed into output, which is what productivity is all about.

The second approach for productivity measurement tackles the problems encountered by labour productivity. As one could imagine, a natural way of dealing with the shortcomings of Equation (1) is by computing a measure of productivity that takes into account all other factors of production, i.e. a measure of multi-factor productivity. Let us define production in firm $i$ at time $t$, $Y_{it}$, as a function $f(\cdot)$ of labour ($L_{it}$), capital ($K_{it}$) and a constant capturing the level of technology, technical efficiency, managerial capacity or any other unobserved components affecting productivity (for a theoretical discussion of this approach see Barro and Sala-i-Martin (1999) and Helpman (2004).

$$Y_{it} = f(A_{it}, K_{it}, L_{it})$$  \hspace{1cm} (2)

Conceptually, the question that we are trying to answer is the following: what is the proportion of output ($Y_{it}$) that is directly attributable to inputs ($K_{it}, L_{it}$) and how much can we attribute to all other aspects of the firm ($A_{it}$)? This last term, $A_{it}$, is called TFP. TFP is better understood as an index number measuring the change in production that is not explained by the change in inputs. There are two major measurement problems to overcome whilst estimating TFP. Firstly, we have to choose a functional form for $f(\cdot)$. This certainly adds a degree of subjectiveness to the estimation imposing a single production function for all firms. Secondly, since $A_{it}$ is actually an unobserved component, a reduced form estimation of Equation (2) without including any kind of control for these unobserved components will result in an over estimation of the importance of inputs, or conversely, the underestimation of TFP.

A traditional functional form of Equation (2) used in, both theoretical and empirical papers is the Cobb-Douglas production function:

$$Y_{it} = A_{it} K_{it}^\alpha L_{it}^\beta$$  \hspace{1cm} (3)

Where $\alpha$ and $\beta$ measure, respectively, capital and labour expenditure shares. Taking logs of (3) and adding a random component:

$$\log(Y_{it}) = A_{it} + \alpha \log(K_{it}) + \beta \log(L_{it}) + e_{it}$$  \hspace{1cm} (4)
As we already mentioned, component $A_{it}$ is unobserved by the econometrician therefore, in practice, we would estimate Equation (4) regressing $\log(Y_{it})$ against $\log(K_{it})$ and $\log(L_{it})$ and assume the residual is composed by elements $A_{it}$ and $e_{it}$. Notice that by assumption $E[e_{it}] = 0$, hence once we have an unbiased estimation of the parameters in equation (3) via OLS (or any other econometric estimation procedure) we are able to get an estimated measure of TFP for each firm as a residual:

$$TPF_{it} = \log(Y_{it}) - (\hat{\alpha} \log(K_{it}) + \hat{\beta} \log(L_{it}))$$

The problem with a TFP estimation like equation (5) is that by not including any control for $A_{it}$ the estimated coefficients, $\hat{\alpha}$ and $\hat{\beta}$, are biased ones. A naïve approach to this problem will specify a fixed-effects estimation within a panel data context. Implicitly such estimation will assume that all those firm-level components giving rise to TFP (adoption of new technology, managerial capacity, etc.) are constant over time. The work of Olley and Pakes (1996) addresses this limitation and, departing from a theoretical model, it proposes an Instrumental Variable (IV) estimation. The main point of the authors is that investment decisions and stocks of capital at the firm level can be used as instruments for the unobserved component of production in equation (3). Investment decisions, according to the authors, can partly capture the entrepreneur’s expected productivity and this can be used as a proxy for the unobserved output component. Using IV estimation, the authors are able to obtain unbiased estimates of the returns to observed inputs ($\hat{\alpha}$ and $\hat{\beta}$ in Equation 5) and therefore their measure of TFP is an unbiased one.

Even if the econometric issues are properly addressed, there is still a further measurement concern that merits a separate discussion. TFP is an index number measuring output over inputs in real terms. In other words, all measures of goods produced and factors of production should be divided by their price before being used in any estimation. Information on prices of every good produced and factor of production used in each firm is rarely available. Therefore, economists have relied on price indexes to deflate outputs and inputs. This is a potential source of miss-measurement and should be carefully considered in any applied work.

There is a special concern regarding price indexes when measuring productivity in the service sector. As mentioned before, a key condition to get an accurate measure of productivity is that both output and inputs are measured in the correct way. Although the goods produced in the manufacturing sector are not completely homogeneous, we assume certain degree of homogeneity in order to use the same price deflator index across output and inputs in different firms. In the case of, say the retail service sector output, the comparability between outputs of two different firms could be misleading (Griffith and Harmgart, 2005). In retail, like in any other business in the service sector, the quality component of the final output is quite high and rather heterogeneous across firms. Quality therefore must be seen as an additional factor of production, which we have to control for. However, this is a concept that is quite difficult to measure.

Finally, it should be noted that despite all the econometric and measurement problems associate with TFP, given the theoretical considerations regarding factor complementarities and marginal productivities, this measure of productivity is the preferred one. However, as it is the case with any other index number, TFP is useful when we are interested in comparative
measurement of productivity. For instance, changes in productivity in two different points in
time or differences in productivity with respect to the mean. Given all the data and
measurement limitations just described, we believe that estimating TFP measures in two
different countries and then comparing the levels of it is not far from being a meaningless
exercise.

2.1.2 Performance measurement

Company performance is assessed in a variety of ways, using measures of productivity, other
financial measures and perceptions of effectiveness. A productivity index can be broadly
defined as the rate of workplace efficiency, and studies have assessed a range of indices at
multiple levels. Organisational performance can be operationalised using either objective or
subjective measures. Hard measures at the firm level include economic or financial metrics
such as return on investment, sales and profitability. Other studies have assessed productivity
in operational terms including individual performance, operating efficiency, wastage indices,
or quality. Soft measures tend to involve senior management’s perceptions of their own
organisation’s productivity as defined in the particular study. These measures are then
usually benchmarked against some appropriate standard, for example, an organisation’s
productivity may be benchmarked against competing firms, to derive meaningful,
contextualised comparisons.

The majority of research has investigated productivity in the manufacturing industry. Indeed,
the manufacture of tangible goods provides more readily available objective and quantifiable
information for investigation. Assessing productivity in the service sector study involves a
related but different set of considerations, for example one author identified excessive
variabilities between service times (Arbós 2002). However, the industry shift from
manufacturing to services can no longer be ignored. More than half of all UK jobs are now in
the service sector, clearly dominating the employment market and therefore the identification
and implementation of more effective management practices in the service sector is likely to
provide the single greatest opportunity for improving the country’s productivity indices based
on this statistic alone.

Some productivity measures are specific to types of company, either by industry or by size of
the organisation. For example, productivity indices of survival are more relevant to small
companies that are relatively financially sensitive, as opposed to multi-national organisations
that have the collateral to withstand greater variations in cash flow. These differences appear
to reflect a real need for context-specificity to enable successful articulation of the
complexities and intricacies of the productivity of companies in different industries and of
different sizes. This also applies to performance measures more broadly, for example,
Silvestro (2002) demonstrated that costs associated with staff turnover will not necessarily
impact upon financial performance unless labour costs are a substantial proportion of total
costs.

Industry-specific measures are typically developed as necessary and ideally empirically
applied within the appropriate work domain, to ensure the resultant productivity measure is
both reliable and valid. Examples include steel production (Ichniowski and Shaw, 1999),
bridge construction (Thomas et al., 2002) and construction (Dunlop and Smith, 2004). This
contingency approach restricts the opportunity for broader comparison and evaluation of
productivity differences and offers some explanation for the discrepancies between published
estimations of productivity ‘gaps’ or differences between the overall productivity of countries.
There is a range of methods that can be applied to derive data suitable for the calculation of various productivity indices. A minority of research has utilised companies’ financial accounting information (Birdi et al., 2006; Bloom and Van Reenen, 2006; Lewis, 2000; Maes et al., 2005; Michie and Sheehan, 2005), or operational or financial organisational performance indicators (Arbós, 2002; Lewis, 2000; Sels et al., 2006).

The majority of the studies reviewed ascertained performance via subjective methods. Questionnaires were employed to determine various productivity indices including operational performance and internal customer satisfaction (Sánchez-Rodríguez et al., 2004); and improvements in organisational performance variables (Hasan and Kerr, 2003). The bulk of questionnaire-based surveys invite respondents to compare their organisation’s performance to that of competitors:

- Financial, marketing and operating performance relative to competitors for the previous fiscal year (Kaynak, 2003).
- Financial performance and labour productivity relative to that of competitors (Department of Trade and Industry, 1999).
- Firm performance relative to that of competitors (Park et al., 2003)
- Organisational and market performance relative to competitors (Chan et al., 2004).

Other researchers have preferred to adopt more direct methods of self-reported data and observations through meetings and interviews with managers, supervisors, and production workers, and 2,594 observations of the production lines spaced at monthly intervals (on average five years worth of data) to calculate actual labour productivity data (Ichniowski and Shaw, 1999); and 215 monthly plant observations to derive the difference between actual and planned standardized hours of labour for plane assembly (Kleiner et al., 2002).

The questionnaire measures assessing productivity focussed on a number of aspects of productivity. Popular names for performance measures are ‘operational’ and ‘financial’ performance, however as discussed earlier these can be operationalised in a number of ways therefore we will investigate the content of measures in more detail. Productivity constructs assessed include:

- Operating performance in terms of inventory management and quality performance (Kaynak, 2003)
- Operational performance measured in terms of the quality of purchased items, on-time delivery, process order cycle time, accuracy, and actual versus target costs (Sánchez-Rodríguez et al., 2004);
- Logarithm of sales per employee (Guthrie, 2001);
- Customer satisfaction operationalised as service quality and measured according to reliability, responsiveness, assurance, empathy, and tangibles (Sánchez-Rodríguez et al., 2004);
- Plant performance derived from annual figures of units per labour hour (adjustments were made for vertical integration (the supply chain), length of the working day, overtime, absenteeism, and for exhaust manufacturers only, product complexity), and quality as indicated by defective parts per million as reported to plants by industry customers (Oliver et al., 1996);
- Improvements in organisational performance: productivity and quality (productivity, efficiency, cost of quality, errors or defects), scheduling and delivery (lead time, timeliness of delivery, vendor relations), Financial results (return on assets, return on sales, return on total quality, market share), and customer satisfaction performance.
(customer satisfaction, employee satisfaction, employee turnover) (Hasan and Kerr, 2003);

- Firm performance relative to competitors’ operating efficiency, quality, service and profitability (Park et al., 2003);

- Management respondents’ ratings of productivity compared to competitors for measures of financial performance and labour productivity published in WERS98 (Bryson et al., 2005; Department of Trade and Industry, 1999).

Chan et al. (2004) determined firm performance using an existing survey tool. Senior executives were required to compare their organisation with similar comparators in the same geographical area (Hong Kong) and make assessments according to various criteria, such as employee relations, product or service quality, profitability, marketing, and so on. In line with the authors’ systems approach, factor analysis extracted two factors that were entitled ‘perceived organisational performance’ and ‘perceived market performance’.

Two studies assessed relative productivity changes over a three-year period. Merino-Diaz de Cerio (2003) collected data from questionnaire-based interviews with knowledgeable employees. Three different indices to quantify improvements were applied. First, cost performance (improvement in percentage of productive hours relative to the total number of hours of direct presence of the workforce). Second, product quality (reduction in the percentage of returned products over sales, reduction in the percentage of defective finished products, reduction in the percentage of defective products in process). Third time-based results (percentage of delivery dates complied with, the time taken from the moment the material is received to the moment where the product is delivered to the customer). All six initial performance measurements were pulled together into a single factor score.

Paul and Anantharaman (2003) measured organizational performance in terms of both operational and financial performance compared to competing organisations during the three-year period, as rated by each company’s CEO or most senior manager available. Operational performance encompasses employee retention, product quality, speed of delivery, employee productivity and operating cost. Financial performance measures comprised of growth in sales, net profit and return on investment for the same three-year period.

Sels et al. (2006) applied three measures of company performance: voluntary turnover (voluntary leavers divided by average staff numbers over one year), labour productivity (value added per member of staff), and financial performance. Financial performance was determined in terms of the share of personnel costs in the value added, net profitability of capital and reserves, the acid test of liquidity and solvency (the degree of auto-financing).

In support of the context-specificity issues outlined above, two studies that took advantage of companies’ published financial accounts (indeed, only publicly quoted companies were included to ensure availability of this data) chose to derive different indices. Maes et al. (2005) selected eight financial ratios due to their relevance to the survival of small businesses: current ratio, acid test, cash flow over equity ratio, profitability of current assets, solvency ratio, degree of self-financing, share of personnel costs in the value added, and the value added per employee. Michie and Sheehan (2005) collated performance indices at the plant level for labour productivity; total sales; and pre-tax profitability. Figures were averaged over a three-year period, to mitigate any simultaneous impact of exogenous variables upon explanatory variable measures during a shorter time window.
2.1.3 Performance measurement in the retail sector

Griffith and Harmgart's (2005) excellent paper stresses the fact that differences in prices of goods and services between two countries (or industries) can reflect differences in market structure rather than differences in quality of output or consumer preferences. For example, if the retail market in country A is more concentrated than in country B, then we would expect that, controlling for everything else, retail productivity will be higher in country A than in country B.

This statement highlights the importance of computing an accurate measurement for both inputs and outputs and to understand the possible reasons behind output-input gaps, which is central to productivity. This is not a problem strictly related to the retail sector, this sort of difficulty can also be present in the manufacturing sector; hence productivity measurement should take into account market structure. The second point raised by Griffith and Harmgart (2005) is the importance played by the technology used in the production process. The authors suggest not only the use of TFP as the productivity measure but they also advocate a deeper insight into the difference in technologies used given the various market conditions (e.g. competitive and investment environment) and firms characteristics (e.g. size, location, etc.).

Reynolds et al. (2005) study is motivated by the US/UK productivity gap in the retail sector. The authors start by asking if there is, indeed, enough evidence showing productivity differentials between US and UK retailers. After revising the available literature on the subject the authors conclude that based upon existing studies, one can conclude that there are enormous measurement problems rather than enough evidence suggesting the existence of a US/UK productivity gap. The authors identify three sources of measurement problems: (1) the difficulty of measuring output accurately in the retail sector (high service/quality component); (2) high level of part time workers in the UK, which are assumed to work half time; and (3) most of the existing studies are based on labour productivity rather than TFP measures.

Given the measurement problems involved in existing studies, the authors proposed two alternatives. The first one consists on what they called retailers approach to measuring productivity which is based on interviews with retail managers to find out what they consider to be important whilst measuring productivity. Out of a sample of 200 UK retailers the authors identify seven areas which almost all retailers identified as a productivity measurement:

- Sales
- Product range
- Service levels
- Availability
- Customer satisfaction (price-value-service-convenience components)
- Employee contribution (often measured in terms of labour turnover)
- Operating and financial performance

Furthermore, the authors identify 21 Key Performance Indicators (KPI) used by many UK retailers (which importance varied by sub-sector and firm size). The KPI can be broadly summarized in the areas of labour, space and capital (Reynolds et al. 2005):

Labour KPIs:
- Labour cost budgets (weekly/monthly) for each store
• Overall labour costs (including as percentage of sales)
• Sales/profit per employee
• Sales/profit per hour worked
• Gross Margin Return on Labour (GMOL)
• Units sold per hour worked
• Till through put (Items per hour going through the checkout till)
• Efficiency ratio (the ratio of hours required to run the store efficiently according to the model, to the actual hours used)
• Staff turnover
• Various customer satisfaction measures

Space KPIs:
• Sales/profit density (sometimes in units per square foot)
• Stock availability (closely relates to and determines space productivity)
• Ratio of selling vs. non-selling space
• Linear density (in an experimental stage for many)
• Trading intensity, or balance of customer traffic, and physical limitations of stores

Capital KPIs:
• ROCE and its variations
• Economic profit or EVA
• Payback period
• DCF-based (Discounted Cash Flow) metrics
• Cost of maintaining the capital base (store base)
• Depreciation as percentage of sales

Based on their interview results, and using data from publicly quoted companies from the UK, the US and France to form a measure of multi-factor productivity the authors find that there is evidence suggesting that labour productivity in the UK is lower than in the US. Here, multi-factor productivity is the productivity of labour, capital and space in a separate way. This is different from the measure of TFP explained in section 3.

2.2 Management Practices Measurement

There is no consensus in the literature regarding how to measure management practices. The only commonality shared by all studies is that management practices are measured in a multidimensional fashion. Due to the inherently intangible nature of management practices, it is challenging to apply objective forms of measurement. Measures are aggregated to facilitate analysis either at the plant, firm, industry or country level.

In the academic literature, these practices are measured using any combination of a variety of scientific methods: self-reported questionnaires, interviews, and observations. Questionnaires and interviews may collect data regarding retrospective or concurrent (or less frequently the prediction of future) management practices. The majority of studies conducting empirical research obtain information by surveying a single knowledgeable individual from each unit of interest, and a minority involve more than one respondent. Less frequently, research studies rely on various unstructured assessment methods, such as observations and analysis of field data collected (Rotab Khan, 2000) and observations alone (Arbós, 2002).
Indeed, the most popular and cost-effective method of collecting empirical data from a large sample is to remotely (usually postal) conduct a questionnaire survey. Another common method is to derive assessments of management practices from structured or semi-structured interviews, whether by telephone or in person. Respondents may be any combination of senior management, HR managers, workplace representatives or the employees themselves. One study interviewed HR managers, operations managers, supervisors, production workers and union representatives (Ichniowski and Shaw, 1999).

Another study requested plant or production managers or, failing that, someone with knowledge of broad organisational and some technical issues (the composite quality management score was calculated across five dimensions: practices relating to the design and development of new products, the production process, links with suppliers, links with customers, and HRM practices (empowerment, training hours per worker, involvement, information sharing) (Merino-Díaz de Cerio, 2003).

Other studies that have assessed management practices solely using questionnaires for remote (largely postal) completion include Total Quality Management (TQM) (Kaynak, 2003), quality practices (Hasan and Kerr, 2003); quality management specifically in purchasing activities (Sánchez-Rodríguez et al., 2004), people management practices (Paul and Anantharaman, 2003); empowerment management practices (Geralis and Terziiovski, 2003); High-Involvement Work Systems (measured in terms of performance-based rewards, alignment, information, involvement, empowerment, teamwork, development, trust, and creativity) (Scotti et al., 2003); High-Involvement Work Practices (assessed against relative use of the following sets of practices: internal promotions, performance (versus service-based) promotions, skill-based pay, group-based (gain-sharing, profit-sharing) pay, employee stock ownership, employee participatory programs, information sharing, attitude surveys, teams, cross-training or cross-utilization, and training to meet future skill requirements) (Guthrie, 2001); High Involvement Work Practices (assessed against relative use of the following sets of practices: internal promotions, performance (versus service-based) promotions, skill-based pay, group-based (gain-sharing, profit-sharing) pay, employee stock ownership, employee participatory programs, information sharing, attitude surveys, teams, cross-training or cross-utilization, and training to meet future skill requirements) (Guthrie, 2001); High Involvement Work Practices (assessed against relative use of the following sets of practices: internal promotions, performance (versus service-based) promotions, skill-based pay, group-based (gain-sharing, profit-sharing) pay, employee stock ownership, employee participatory programs, information sharing, attitude surveys, teams, cross-training or cross-utilization, and training to meet future skill requirements) (Guthrie, 2001); Human Resource Management (HRM) intensity (Sels et al., 2006) and operational practices (Maes et al., 2005).

Indeed, sometimes the method of data collection needs to be tailored to cultural requirements. A study assessing management practices in identified Japanese subsidiaries in both the USA and Russia made a special effort to set actual interview times with organisations in Russia only to talk respondents through the questionnaire (Park et al., 2003). This was necessary because Russian organizations are traditionally very protective of company information, and therefore require direct assurances to be willing to share this with people external to the organisation. Oliver et al. (1996) used a lengthy survey tool involving 880 data fields, which took on average four days of management time to complete. The research team conducted an initial visit to each organisation to explain how to complete the survey, and then made a second visit (up to one whole day) to review the information provided. These meetings were conducted to enhance both response rate and the quality of the data collected.

When relying upon self-reported data, it is important to set out a consistent frame of reference that all respondents can both understand and relate their answers to. When an organisation operated from more than one site, Michie and Sheehan (2005) asked the interviewee to respond to questions in terms of the site that most typified the organisation’s HR practices. Some studies require respondents to benchmark the quality of their employer’s management practices against those of their competitors (Park et al., 2003), whereas others require the percentage of staff covered by each particular management practice (Chan et al., 2004), or the
presence of particular practices and are asked to estimate the ‘richness of adoption’ (Lewis, 2000).

Some researchers have adopted precise methods of assessment to facilitate more sensitive analyses. Sels et al. (2006) assessed defined, complementary sets of HRM practices to yield ratings for different HRM domains. This technique was adopted to record more gradual differences between organisations, and allowed the authors to investigate the relative intensity of HRM practices between organisations. Another measurement technique considered to be more sensitive was applied by Kaynak (2003), whereby respondents rated their organisation’s actual adoption of TQM practices. Usage was indicated by marking a point along a continuous 100mm scale.

Considering real-time interviews, one study involved interviewers who were all specially trained graduate students with at least five years’ business experience, and therefore all well-qualified to assign ratings for each company’s management practices (Bloom and Van Reenen, 2006). This arrangement facilitates the removal of interviewer fixed effects, which is otherwise a source of error when a different individual responds for each organization. As the interviewers judged the ratings, a pivotal ‘funnelling’ telephone interviewing technique was applied using open questions. The novel survey instrument targeted manufacturing practices, and questions were grouped between operations, monitoring, targets and incentives. Operational data encompassed the introduction, rationale for, and documentation of lean manufacturing process improvements. The monitoring section concerned employees: tracking, reviewing and managing consequences of performance. Different characteristics of targets are incorporated, namely: type, realism, transparency, and their range and inter-dependency. Incentives comprised of promotion criteria, addressing bad performers, pay and bonuses, all with the ideal of strongly rewarding both effort and ability. Data derived from the main survey of plant managers was cross-validated with information collected from two other surveys: one with the HR department, and a follow-up survey with the CEO or his office.

Only one study investigated a truly longitudinal data set. Birdi et al. (2006) matched data from three separate telephone interviews with each firm conducted over a seven-year period. The same interview structure was followed at each time point: a definition of each practice was read out, and the respondent asked to confirm whether or not it had been introduced by their organisation and if so in which year.

Bloom and Van Reenen (2006) applied a ‘funnelling’ telephone interviewing technique using open questions, and the main survey of plant managers was supported with information collected from two other surveys: one with the HR department, and a follow-up survey with the CEO or his office. Birdi et al. (2006) matched data from three separate telephone interviews conducted over a seven-year period to create a longitudinal data set. The same interview structure was followed at each time point: a definition of each practice was read out, and the respondent asked to confirm whether or not it had been introduced by their organisation and if so in which year.

Bryson et al. (2005) measured HIM practices using data sourced from WERS98 (Department of Trade and Industry 1999). Nine management practices were grouped into three meaningful sets of indicators: task practices (functional flexibility, team working and problem-solving groups), individual supports (information disclosure, team briefing, training in problem-
solving or communication), and organisational supports (job security guarantees, broad-based financial participation schemes, emphasis on internal promotion).

There are already established questionnaire measures of particular management practices, and most studies chose to utilise these in their original or a modified format. Typically only in the absence of a suitable existing tool do researchers choose to develop their own instrument. There are various questionnaire formats and they are not specific to particular management practices, the following examples are for illustrative purposes only.

Sels et al. (2006) measured complementary sets (of three) HRM practices to measure each HRM domain. This technique was adopted to allow for more gradual differences between organisations, and allowed the authors to investigate the relative intensity of HRM practices between organisations. The domains that have been surveyed are training, selection, compensation, careers, performance management, and participation.

To ascertain the actual adoption of TQM practices, Kaynak (2003) asked respondents to rate firm usage of particular practices. Usage was indicated by marking a point along a continuous 100mm scale for each of the following types of practices: management leadership, training, employee relations, quality data and reporting, supplier quality management, product/service design, and process management. Oliver et al. (1996) measured the use of lean production management practices with quantitative survey indicators. In particular, operational ‘leaness’ was determined by counting the hours of inventory at specific stages of production, and amount of effort required for reworking repair versus the production of new goods. However, these two measures may be better construed as indicators of the effectiveness of lean management practices. Further items measured the presence or absence of team responsibility, continuous improvement schemes, high-commitment human resource policies and supply chain management.

Where an organisation operated from more than one site, Michie and Sheehan (2005) asked the interviewee to respond to questions in terms of the site that most typified the organisation’s activities. Eight areas of HR practices were measured: training, formal appraisals, recruitment and selection, internal career opportunities, employee voice (participation) and consultation, job design/internal employee flexibility, employment security and performance-based pay.

Some measures require respondents to benchmark their employer’s management practices against those of their competitors. For example, Park et al. (2003) asked questionnaire respondents to rate their employing organisation against that of its competitors along a number of Likert-type items in the following categories: employee skill, attitudes, motivation, and HR systems (for example, our organisation places importance on training, and employee input and suggestion are highly encouraged).

Lewis (2000) applied an integrated set of research methods and measures. Management completed a self-reported questionnaire based on a series of indices describing lean production tools and techniques. Respondents were asked to indicate whether these were present in their firm, and estimate the ‘richness of adoption’. This criterion was used to select three firms for case studies to explore the author’s exploratory propositions. From these firms, data was collected from interviews with managers and team leaders using semi-structured question sets, in addition to secondary data such as company accounts and employee surveys. Kosonen and Buhanist (1995) also employed a combination of methods
and measures to evaluate both technical and human components of the case study organisation. Data was collected from team and individual interviews, observations and questionnaires.

Bloom and Van Reenen (2006) devised a novel survey instrument to assess management practices. Questions are grouped into four areas: operations (3 questions), monitoring (5 questions), targets (5 questions) and incentives (5 questions). Operational data encompassed the introduction, rationale for, and documentation of lean manufacturing process improvements. The monitoring section concerned employees: tracking, reviewing (for example using regular appraisals and job plans) and managing consequences of performance (such as appropriate administration of sanctions and rewards). Different characteristics of targets are incorporated, namely: type (purely financial or more holistic), realism (stretching, unrealistic, or non-binding), transparency (simple or complex), and their range and inter-dependency (such as how consistently they are followed throughout the organisation). Incentives are comprised of promotion criteria, addressing bad performers, pay and bonuses, with the ideal of strongly rewarding both effort and ability. The survey tool is groundbreaking because it offers a way of assessing and incorporating management practices at a grassroots level.

Management practices are multi-dimensional constructs and therefore some researchers have adopted a systems approach. A smaller number of promising studies have assessed intervening variables regarding processes linking practices to performance. Paul and Anantharaman (2003) assessed the mediating role of HRM practices through the following processes: employee competence, teamwork, organizational commitment and customer satisfaction.

Chan et al. (2004) measured High Performance Human Resource Practices (HPHRP) using an existing instrument, modified for use in Hong Kong. HR Managers were required to rate each specified practice in terms of the percentage of staff at their organisation who were covered by it. Three other questions (promotional criteria, qualified applicants for job vacancy, and training hours) were measured in terms in a forced-choice or numerical response format. Factor analysis resulted in two factors, reflecting a systems approach to HR practices, labelled ‘employee skills and organisational structure’ and ‘employee motivation and communication’.

### 2.3 Empirical frameworks and level of analysis

#### 2.3.1 Empirical frameworks

Regarding the impact of Just In Time (JIT) and TQM on productivity/performance, some papers have estimated the beneficial effect of management practices in terms of costs reduction by using a CES-TL cost model and a Full Information Maximum Likelihood estimator (Brox and Fader, 2002) or an OLS estimator (Brox and Fader, 1996; Brox and Fader, 1997; Callen et al., 2000). Other strands of studies have either estimated the impact of those management practices on proxies of performance by using an OLS estimator (Lawrence and Hottenstein, 1995; Huselid, 1995; Black and Lynch, 1996; Koch et al., 1996; Caroli and Van Reenen, 2001) or they have estimated the technical efficiency of a firm by means of a Stochastic frontier methodology (Kaynak and Pagan, 2003; Callen et al., 2005). Other studies simply assume that management practices increase productivity and performance but they do not rigorously prove it. They tend to focus on cost/productivity differences that exist between groups of adopters versus non-adopters exposing their results to problems of unobserved
heterogeneity or lack of control of lurking factors – (see for example, Dorgan et al., 2006, Sale et al., 2003).

That is not for example, the case of (Patterson et al., 2004) that estimate the impact of a cluster of management practices on performance by more robust means of multiple regression analysis. The impact of a cluster of management practices on productivity has also been estimated by using Logit and Probit estimators (Ichniowski et al., 1995), bivariate analyses of correlation (Ichniowski et al., 1997), or even more robust empirical frameworks such as IV, WG (Within Groups) and GMM estimators of panel data models (Bloom and Van Reenen, 2006). Robust techniques have been used also in the estimation of an inverse causality relationship between management practices and productivity. This relationship is estimated by using both OLS, Probit and a DPD (Dynamic Panel Data Model) estimator (Nickell et al., 2001).

Regarding the impact of IT or ICT on productivity, the empirical methodologies used have been various, as for example, descriptive statistical analysis of the characteristics of different groups of adopting firms (Swamidass and Winch, 2002; Dorgan and Dowdy, 2004), WG or GMM estimators (Lehr and Lichtenberg, 1999; Black and Lynch, 2001), OLS estimator (Licht and Moch, 1999; Wolf, 1999; Gera et al., 1999; Ramirez et al., 2001; Basu et al., 2003) and even more robust non-parametric estimators (Kaiser and Bertschek, 2004).

Nevertheless, one of the major shortcomings of the studies in this area is the lack of sufficiently long longitudinal data at firm (or plant) level. This has, indeed, severely constrained the possibility to test in a robust way the presence of complementarities between two innovations; the direction of their adoption sequence and the lagged impact of the adoption decision upon firm performance.

Several studies use cross-sectional analysis. However, the latter may be biased because of endogeneity problems, i.e. the presence of unobserved firm characteristics that are time invariant. In addition, estimated coefficients may be affected by measurement errors. Conversely, using a panel data may provide several advantages. First, with longitudinal data, it is possible to address the problem of “unobserved heterogeneity” using a Within Estimator to control for time-invariant unobserved characteristics. Second, longitudinal data on inputs may allow us to examine the lag structure such as how the accumulation of management practices over time, within a firm, affects the current productivity. Third, with longitudinal data by using a Generalised-Method-of-Moments it is possible to address some of the issues associated to the measurement errors.

2.3.2 Level of analysis

With respect to the level of analysis, the vast majority of papers investigates the link between management practices – as JIT/TQM, HRMPs and ICT - and productivity at firm-level or industry-level, whilst fewer papers have focused their analysis at plant-level (Lawrence et al., 1995; Callen et al., 2000; Callen et al., 2005; Swamidass et al., 2002; Black and Lynch 2001) or establishment-level (Black and Lynch, 1996; Ichniowski et al., 1997; Sale et al., 2003; Caroli et al., 2001).

It is reasonable to think that results from these studies may be arguable to some extent. This is due to the wide range of subjective definitions and measurements of both management practices and performance or productivity, which makes comparisons difficult.
2.3.3 Inverse Causality Relationship between Management Practices and Productivity/Performance

Some recent studies have investigated the causal relationship between management practices and productivity. Particularly interesting is the paper of (Nickell et al., 2001) not only because it uses a panel data (rather than cross-sectional data) of both firms and plants, but also because it reports evidence about an inverse causal relationship between firm performance and introduction of a cluster of management practices. In particular, from a theoretical ground it develops a dynamic model in which the firm chooses employment, total hours (hours spent on production and hours spent to improve the organization, product quality etc.) and their allocation in order to maximize profit subject to the evolution of a productivity factor, which is function of the hours per week spent by workers. From an empirical viewpoint it investigates not only the relationship between management innovations/practices and changes in the real situation of the firm (performance) by using a DPD estimator, but also whether past changes in performance of plants explain the adoption of future management innovations, by using both the OLS, Probit and Ordered Probit estimator.

Management practices or innovations are defined here not as technological innovations but as improvements in the way things are done and they are measured in a number of spheres such as: reductions in restrictive practices by employees, introduction of new technologies, changes in the organizational structure, increases in decentralization, adoption of new human resources management practices, changes in the industrial relations and initiation of new JIT practices. Performance is measured by the real profit per employee that, to some extent, captures both prices and productivity changes.

Results of the analysis show that, consistently with the predictions of the theoretical model, a worsening of the firm performance – due to a reduction in productivity or price of output - will lead to the introduction of management innovations both for firms and plants, and that firms with higher debt burden (under financial pressure) are less likely to introduce management innovations.

Two reasons are given for the fact that it is when times are bad and business slack that firms introduce new management practices, reorganize production methods and so on. First, when demand is slack and profitability is low, both managers and workers have more time to devote to organizational issues. Second, bad times mean a higher probability of bankruptcy and an increased threat to jobs. Almost inevitably, the response of both managers and employees will be to try and lower this threat by reducing the chances of the firm going bankrupt. One way of doing this is to set about improving productivity. However, it is worth noticing that this study does not take into account that, in presence of imperfect capital markets, if the firm is under financial pressure, the ability of firm to reorganize or introduce new technologies may be reduced. Also, the evidence about the fact that, on the one hand, management practices affect performance and on the other hand firm performance may affect management practices might suggest that empirical results of the extant literature be misleading because likely biased for not controlled endogeneity.
3. Management Practises

3.1 What Management Practices Are Studied?

From high-level country measures of aggregate productivity, down to the productivity of a single branch of a firm or individual employee, many factors may exert an influence. Factors may include the amount of financial investment available, stocks of human and physical capital, the institutional environment and the implementation of management practices, amongst others.

From an economist’s point of view, managers’ behaviour will always be optimum given the information available in the market; hence assuming a perfect flow of information, management practices will be considered state variables rather than determinants of productivity. However, in a world of imperfect and asymmetric information, the role played by managers in enhancing productivity may account for a significant proportion of between-firm variation.

Studies that investigate the link between management practices and productivity have assessed the impact of an individual management practice in isolation, the effects of joint adoption or the impact of clusters or systems of complementary HRM practices. Practices can be meaningfully separated out into operational and HRM practices. Operational management practices include ICT, TQM, and lean production. HRM practices focus on people management; in particular the recruitment, development and management of employees (Wood and Wall, 2002). Despite the existence of a multitude of contrasting conceptualisations of these practices, there is a lot of common ground between them (Wall and Wood, 2005). Typical HRM practices involve training and development, empowerment, and teamwork.

The melange of management practices studied can be attributed to not only the contrasting onuses of different theoretical perspectives, but also more importantly a popular undercurrent of context-specific research. Contingency theories predominate in the HRM literature in particular Wall and Wood (2005) suggesting it is unlikely that there exists a ‘one size fits all’ set of productivity-enhancing management principles or practices. Edwards et al. (2004) builds upon this contingency approach, stating that the success of management practices are firm-specific and these are affected by the prevailing institutional environment.

Researchers have investigated the impact of various management practices on productivity: TQM (in terms of management leadership, training, employee relations, quality data and reporting, supplier quality management, product/service design, and process management) (Kaynak, 2003); quality practices (Hasan and Kerr, 2003; Sánchez-Rodriguez et al., 2004); HR, operationalised by Ichniowski et al., (1997) in terms of: incentive pay schemes, recruiting, teamwork, employee’s security, flexible job assignment, skills training, and communication and labour relation practices. Michie and Sheehan (2005) investigated eight areas of HR practices: training, formal appraisals, recruitment and selection, internal career opportunities, employee voice (participation) and consultation, job design/internal employee flexibility, employment security and performance-based pay) or people management (Paul and Anantharaman, 2003) measured eight practices: selection, induction, training, job design, work environment, performance appraisal, compensation, career development and incentives) practices (Ichniowski et al., 1997; Michie and Sheehan, 2005; Park et al., 2003; Paul and Anantharaman, 2003); empowerment management practices (Geralis and Terziovski, 2003); High-Involvement Work Systems (measured in terms of performance-based rewards,
alignment, information, involvement, empowerment, teamwork, development, trust, and creativity) (Scotti et al., 2003); High-Involvement Work Practices (assessed against relative use of the following sets of practices: internal promotions, performance (versus service-based) promotions, skill-based pay, group-based (gain-sharing, profit-sharing) pay, employee stock ownership, employee participatory programs, information sharing, attitude surveys, teams, cross-training or cross-utilization, and training to meet future skill requirements) (Guthrie, 2001); HRM intensity (Sels et al., 2006); managerial characteristics (determined in terms of managerial ownership, incentive schemes, organisational flexibility, formality in management structure, training, collaboration and industry characteristics) (Cosh et al., 2005); operational practices (Maes et al., 2005); and a mixture (Waterson et al., 1999) investigated the impact of a mixture of HRM and operational management practices, namely: business process re-engineering, supply-chain partnering, outsourcing, learning culture, empowerment, team-based working, total productive maintenance, concurrent engineering, integrated computer-based technology, manufacturing cells, just in time production and TQM) of HRM and operational management practices (Birdi et al., 2006; Bloom and Van Reenen, 2006; Waterson et al., 1999).

3.2 ICT and Productivity/Performance: Individual and Joint Adoption.

Economists have long speculated on why there are differences in the productivity performance between firms. Although a strand of economic research has tried to overcome the problem by looking for better measures of inputs, a consistent part of the literature has tried to see how much of the residual can be accounted for by explicit measures of technology as Research and Development of Information Technologies or Computerization. Even if a substantial unexplained productivity differential still remains, IT is certainly a part of the story.

IT usage has permeated virtually every sector of modern economies, and for decades the world IT sector experienced significant growth. The swift development of ICT, as well as the declining prices for its use, have considerable enhanced the diffusion of ICT during the recent years.

Consequently, the impact of ICT on productivity has become the main topic of discussion in economics and management sciences. Although the research on that impact usually starts with the assumption that “computers enhance productivity”, the related empirical evidence is mixed. As some studies at industry-level have failed to detect a positive contribution of ICT on productivity, other studies, using firm-level data, have found empirical evidence for a positive productivity effect or no productivity effect of ICT.

From a firm level perspective, the paper of Swamidass and Winch (2002) provides evidence about a positive but different extent of impact of manufacturing technology innovations (ICT) on productivity (sales per employee) between the US and UK manufacturing plants, over the period 1997-1998. Descriptive statistical analysis is performed to compare the extent of use of technologies between the UK and the US plants. Evidence from this study shows that, the investment in ICT has been similar between the UK and US even if the higher levels of computerization for the US manufacturing plants appear associated to higher levels of productivity and Returns On Investments (ROI) with respect to the UK plants.

Another study supporting the evidence of a positive impact of IT on firm productivity is Lehr and Lichtenberg (1999). By using a large sample of US firms, over all sectors and over the
period 1977-1993, and a Within Group estimator, it estimates the impact of IT on firms’ productivity growth rate (proxied by sales growth rate), when IT is measured as the share of computer capital in the total capital stock. Empirical results show that computerization positively affect the productivity growth rate of firms.

By contrast, the paper of Licht and Moch (1999) provides evidence of no effects of IT on productivity. By using a large sample of German manufacturing and services establishments, over 1996, and a Cobb-Douglas production function, it estimates, by means of an OLS estimator, the impact of IT on labour productivity. Results show that, although a large number of firms claim to have realized productivity gains over the period, IT investment does not show effects on the labour productivity.

Mixed evidence about the sign of the impact of IT on productivity is also reported at industry-level.

Some studies (Wolf, 1999) show that computerization – defined as office, computing and accounting equipment for employee - has a negative impact on total factor productivity of the service sector. This result seems to support the common view that in the services sector computers make things to work only a little more efficiently or the view that quality aspects of technical change are hard to assess in this sector.

Gera et al. (1999) that estimates the impact of IT investments on labour productivity growth of Canadian and US manufacturing and services industries, over the period 1971-93, finds that IT investments positively affect the productivity growth and IT investment is much more important than non IT investment in affecting labour productivity growth both for Canadian and US sectors.

In an attempt to investigate whether the slowdown UK productivity growth, both as TFP and labour productivity, in the second half of 1990s, could be due to the low level of investments in ICT, Basu et al. (2003) estimates, by means of an OLS estimator, the impact of ICT – as the average share-weighted computer and software capital growth rate – on the TFP growth rate – as the average TFP growth rate over the period 1995-2000. The picture emerging from the empirical results is that in the late 1990s the TFP growth was strongly and positively associated with the growth of ICT capital services.

Similarly, O’Mahony and Van Ark (2005) report evidence about a positive impact of the adoption and diffusion of ICT on productivity growth in the UK retail trade sector, in comparison with the US, Germany and France over the period 1995-2000.

In particular, the impact of investment in ICT on labour productivity growth is estimated by using the “Growth Accounting Methodology”, which is a method to decompose the output growth into contributions of various factor inputs, weighted by their shares in the value of output (more specifically, output growth = inputs growth (volume of L and K, and quality of L and K) – TFP - all other factors affecting output growth). This method, hence, allows measuring the labour productivity growth and the contributions to it given by capital (as ICT and non-ICT capital), quality of labour and TFP.

The papers by Stiroh (2002), O’Mahony and Robinson (2003), and Vijseelaar and Albers (2004) use data at industrial level to estimate the relationship between labour productivity/TFP and ICT. Stiroh (2002) uses the difference-in-difference estimator to

The main result coming from these three papers is that, although ICT has a positive correlation with TFP, there is not enough evidence supporting the new economy hypothesis (whether the increase in ICT investment is the reason behind the rise in US productivity during the second half of the 1990s: this hypothesis is what has been called the new economy, whereby assuming a strong spill-over effects of ICT there is a rise in aggregate productivity). In other words, there are few spill-over effects of ICT into the rest of the economy; therefore, ICT cannot account for the observed increase in productivity in the US during the second half of the 1990s. Similarly, the survey papers by Ignazio (2000) and Pilat (2004) support the same findings.

Basu et al. (2003) challenges the conventional view that ICT has no spill-over effects and therefore cannot contribute to explaining the US/UK productivity gap. The authors’ main idea is that investment in ICT has a lagged effect upon TFP. Moreover, contemporaneous investment in ICT can have a negative effect upon TFP. If this hypothesis is true, then the main implication of all studies trying to identify ICT as the main source of productivity differential between the US and the UK is that ICT cannot explain such a gap. Taking data for the whole US economy at industrial level, Basu et al. (2003) find that growth in ICT between 1980 and 1990 has had a positive effect upon TFP growth between the years 1995 and 2000. Conversely, growth in ICT between 1995 and 2000 has been negatively correlated with growth in TFP during the same period. For the UK the evidence was not conclusive: lagged ICT growth has not affected the present TFP growth, although present ICT growth was negatively related with TFP growth. Given that the UK investment in ICT during the 1980s was lower than the ICT investment in the US, the lagged effect of ICT growth upon TFP growth could - at least partly - explain the US/UK productivity gap.

Often optimistic European policy makers claim that boosting IT spending may contribute significantly in closing the UK productivity gap with the US, while a strand of academic research argues that a crucial role is also played by good management practices which may mediate the impact of IT on productivity. In fact, whilst a strand of literature that has examined the influence of the IT on firm’s performance argues that the adoption of a IT-based process innovation by itself leads to productivity gains, more recent studies argue that a IT-based process of innovation by itself does not lead to productivity gains if it is not accompanied by associated innovations in the production organization sphere, new customer and supplier relationships and new product design (see for example Bresnahan et al. (2002) or Battisti et al. (2005). In order to give a support for the presence of complementarities Caroli and Van Reenen (2001) show that plants that have introduced computer based equipment are more likely to be involved in organisational changes, with positive effects on productivity. Battisti et al. (2005) by using a panel data set of Italian plants, find that there exist complementarities between ICT related innovations and work practices innovations. However, they could not determine the direction of the adoption sequence. They also find that due to complementarities joint adoption increases the benefits from joint adoption over and above those derived from individual adoption. Bresnahan et al. (2002) find that skilled labour is complementary with a cluster of three distinct changes at the firm level: information technology, new work organization and new products and services.
Other studies have focused not only on the existence of complementarities between HRM practices and ICT but also on their impact on productivity or performance (some papers, see for example Shadur et al. (1995) limit their analysis to the observation that IT high productive companies are characterized by efficient HR management practices). For instance, Ramirez et al. (2001) analyses both, the impact of three management practices – Employee involvement work practices, TQM and Reengineering – on the added value of output, and the impact of these three management practices on the productivity of IT investments. In particular, by using a sample of UK manufacturing and services firms in 1996 and a Cobb-Douglas production function as a function of IT, capital, labour, management practices and an interaction term of management practices with IT, it estimates – by means of an OLS estimator – both the direct impact of management practices on output, and the indirect impact of those practices through IT (that is the IT productivity contribution). An interesting result emerging from the estimated parameters is that the implementation of management practices positively impacts the productivity of IT investments. Also, this impact results higher for manufacturing firms, rather than for services firms, which invest more in management practices than services firms.

Another study that takes into account the crucial role of management practices for IT investments and its impact on productivity is Dorgan and Dowdy (2004). This study supports the view that IT expenditures have little impact on productivity, unless they are accompanied by first-rate management practices. In particular, Dorgan and Dowdy (2006) describe and compare financial characteristics of 100 manufacturing firms across France, Germany, UK and US over the period 1994-2002, and rate these firms on a scale of 0 to 5 according to their degree of use of three management practices – lean manufacturing, performance management and talent management. According to the results of their statistical analysis they conclude that management practices may have a positive impact on firm performance, since they observe that a one-point improvement in the scale is correlated with 5 percentage point increases in the company ROCE – as proxy for firm performance – and it is correlated with 25 percent increases in the company’s TFP. Also, the impact of IT investment appears modest and companies with more powerful IT do not perform better financially. However, management practices, they argue, can increase the impact of IT investment on productivity; in fact, companies with increased computing power and improved management practices achieve 20 percent higher productivity.

Although it seems reasonable that ICT has an indirect effect on labour productivity by enabling firm’s reorganization of workplaces, researchers only recently became more concerned with the joint effects of workplace organization and ICT on labour productivity (Black and Lynch, 2001; Kaiser and Bertschek, 2004; among others).

The strand of literature that analyses the relationship between investment in ICT, workplace reorganization and labour productivity takes the view that ICT and workplace reorganization positively affect labour productivity. For example, Black and Lynch (2001), using a sample of 600 US manufacturing firms, provide empirical support to this view. In particular, they examine the impact of workplace practices, information technology and human capital investments on productivity. They estimate a Cobb-Douglas production function with both cross section and panel data covering the period of 1987-1993, using both Within Group and GMM estimators. Results of their econometric analysis show that it is not whether an employer adopts a particular work practice but rather how that work practice is actually implemented within the establishment that is associated with higher productivity. Moreover,
plant productivity is higher in businesses with more-educated workers or greater computer usage by non-managerial employees (Leseure et al., 2004).

Furthermore, there is across these studies some evidence about a potential reverse causality between labour productivity and workplace reorganization. Labour productivity can itself affect workplace reorganization. However, Kaiser and Bertschek (2004) seem to overcome this critique.

Kaiser and Bertschek (2004) analyse the relationship between investment in ICT, non-ICT, labour productivity, and workplace reorganization. This study presents two main merits. First, it takes into account the potential simultaneity between labour productivity and firms’ decisions to reorganize workplaces by estimating an endogenous switching regression model for a sample of 411 firms from the German business-related services sector. Second, it allows for complementarities between the input factors and workplace reorganization.

In particular, by assuming that firms reorganize the workplace if the productivity gains arising from workplace reorganization exceed the related reorganization costs, a model is developed and estimated for labour productivity and the firms’ decision to reorganize workplaces that allows workplace reorganization to affect any parameter of the labour productivity equation.

The estimation results show that changes in human resource practices do not significantly affect firms’ output elasticities with respect to ICT, non-ICT capital, and labour. However, the point estimates with respect to non-ICT investment and labour tend to be larger if workplace reorganization takes place. Therefore, a Kernel density-estimation technique is applied to show that for firms with organizational changes, the entire labour productivity distribution shifts significantly out to the right if workplace reorganization takes place, indicating that workplace reorganization leads to an increase in labour productivity that is attributable to complementarities between the various inputs factors and workplace reorganization. By contrast, firms without organizational changes would not have realized significant productivity gains if they had reorganized workplaces.

In sum, concerns about an “information technology productivity paradox” were raised in the late 1980s. The productivity paradox of IT was that, despite enormous improvements in the underlying technology, the benefits of IT spending have not been found in aggregate output statistics. One explanation was that IT spending may lead to increases in product quality or variety which tend to be overlooked in aggregate output statistics, even if they increase sales at the firm level. Nowadays, the understanding of the relationship between information technology and economic performance has improved. Overall, firm level studies have suggested that, rather than being paradoxically unproductive, IT when combined with new practices has an impact on economic growth (Brynjolfsson et al., 2000).

3.3 JIT/TQM and Productivity/Performance: Individual and Joint Adoption

Within the literature analysing the impact of the adoption of individual management practices on productivity, some studies, rather than directly focusing on the impact of the practice on productivity, analyse the beneficial effect of its adoption in term of costs reduction. This implicitly assumes, but does not prove, that firms adopting management practices are on new more efficient production functions. Some papers, for example, have developed mathematical models to find the positive impact of the optimum Just-in-Time on the firm’s expected total
cost by challenging the view that inventories are of no value and should be totally eliminated (Salameh and Gattha, 2001).

From an empirical perspective, studies have quantified the beneficial effect of individual management practices on a firm’s costs or productivity by using several methods. In particular, by means of parametric estimation methods, many studies have analysed either the impact of one management practice upon cost function (by estimating for firms classified as management practices users and not users a CES-TL cost model) or they have simply estimated the relationship between management practices and proxies for productivity/profitability. For example, Brox and Fader (1996) as well as Brox and Fader (1997) explore the cost function differences between samples of JIT and non-JIT user firms by estimating – via OLS - a generalized CES-translog cost model. The empirical findings support the view that JIT management practices enhance productivity and cost efficiency. JIT practices are defined in this context as JIT/TQM practices, namely: Kanban, Integrated product design, Integrated supplier network, Plan to reduce set-up time, Quality circles, Focused factory, Preventive maintenance programs, Line balancing, Education about JIT, Level schedules, Stable cycle rates, Market-paced final assembly, Group technology, Program to improve quality (product), Program to improve quality (process), Fast inventory transportation system, Flexibility of worker's skills. Moreover, whilst productivity is measured as labour productivity (output/labour), profitability or performance is measured as profit to investment.

One drawback of this analysis is that it does not effectively estimate the impact of the single management practice on productivity/performance, but it simply analyses and compares financial characteristics of JIT and non-JIT firms.

Similarly, Brox and Fader (2002), classify firms as non–JIT or JIT users not only according to a self-declaration but also according to the management strategies designated by the survey to capture the extent of JIT use. Then, they estimate for each sub sample of firms a CES-TL cost model, by the method of Full Information Maximum Likelihood. Also, they provide a statistical analysis of any cost/productivity difference that exists between groups of manufacturing firms that have adopted the JIT and those that have not. The main finding of this analysis is that firms following the set of JIT management strategies are more profitable than non-JIT firms in the same industry: there are indeed significant gross profit differences between JIT and non-JIT users. Moreover, the cost elasticity with respect to output is lower for the JIT group of firms indicating that they are better able to capture economies of capacity utilization.

Unlike previous works, there are studies that provide a more sophisticated analysis of the impact of management practices on productivity. Callen et al., (2000), for example, by measuring profitability (rather than productivity) as profit margin (operating profits divided by sales revenues) and contribution margin ratio (contribution margin divided by sales revenues), in order to investigate whether JIT is associated with greater plant productivity, improved quality of process and product, lower costs and higher profits, runs an OLS regression analysis on manufacturing plants data. Even in this case, results show that JIT plants are significantly more profitable than non-JIT plants. Similar results are found by Lawrence and Hottenstein (1995) that, by using proxies of performance (quality, lead time, productivity and customer’s services) and proxies of JIT management practices (the extent of employees participation, suppliers participation and management commitment) estimate by
means of an OLS the relationship between JIT practices and performance for a sample of Mexican plants affiliated with US companies.

By means of simple statistical analysis of correlation, Dorgan et al. (2006) find a positive correlation between improvements in the quality of management practices and improvements in TFP and performance for a sample of 700 midsize manufacturing companies across France, Germany, UK and US. Management practices are, here, measured as lean production methods, techniques for setting targets and JIT production: managers are asked to rank on a scale of 1 to 5 improvements in the quality of management practices; whilst, performance is measured by a wide set of financial variables – such as market share, sales growth, market capitalization and ROCE (proxy for the Tobin’s Q) – and productivity as TFP. The latter is an efficiency measure that captures the impact of all the elements that contribute to a company’s output growth, but are not explicitly stated as production factors. It is, in other words, a grab bag for the unexplored elements – such as technology, public infrastructures or management practices – that affect productivity. This casts doubts on the reliability of results coming from the extant literature that typically uses to estimate the impact of management practices on TFP. In fact, as TFP includes the effects of management practices, we should expect those estimates be biased for endogeneity.

A strand of research has even studied the impact of an individual management practice on productivity by using a Stochastic Frontier Methodology. This method assumes that, for a given technology, it is possible to relate inputs to output through a production function, and that this function has two error terms: a symmetric random error accounting for unobservable effects and a non-symmetric random error term accounting for productive inefficiency. Given the Translog Production specification with two inputs (L and K), the frontier equation for the firm is estimated with a two step procedure: in the first step the stochastic frontier is estimated to obtain technical inefficiency estimates; in the second step these estimates are regressed on determinants of inefficiency.

Given that a firm is said to be technically inefficient if it is not able to reach maximum output given its available resources of technology, Kaynak and Pagan (2003) capture JIT related sources of technical inefficiency by formulating a stochastic frontier model where the parameters of a translog production function are estimated simultaneously with the technical efficiency effects. More specifically, it is estimated the impact of four JIT characteristics on the way in which firms combine resources to produce a given level of output. The testable hypotheses are: a) the higher the cooperation of suppliers in JIT implementation the higher the productive technical efficiency achieved by firms, b) the higher quality materials suppliers provide to firms the higher the productive technical efficiency achieved by firms, c) the more transportation activities of firms are aligned with JIT implementation the higher the productive technical efficiency achieved by firms, d) the higher the commitment of top-management to making JIT a priority for the whole organization the higher the productive technical efficiency achieved by firms.

The empirical results suggest that characteristics internal to the organization, such as top management commitment to implementing JIT, are related to higher productive efficiency. External characteristics, such as supplier value-added, or transportation issues, do not appear to be associated with increasing productive efficiency. Also, by means of a statistical analysis, it is found that the degree of implementation of JIT is significantly related to each performance factor: financial and market performance, time-based quality performance, and inventory management performance.
Kaynak (2003) is instead a paper reporting evidence about the impact of TQM on firm’s performance. Indeed, by using a combined sample of manufacturing and service firms, it shows a positive relationship between the extent to which companies implement TQM and firm performance. Three TQM practices (specifically: process management, supplier quality management, and product or service design) exert a direct effect on operating performance, and other TQM practices indirectly affect operating performance via those three practices. Operating performance mediates a positive effect of TQM practices on financial performance.

A recent contribution to the literature is given by Callen et al. (2005). It analyses the interaction among performance outcomes, investment in JIT management practices, and productivity measurement at the plant level. On the one hand, it estimates a stochastic frontier production function (as function of labour, capital, fuel, and JIT technological index) and provides an analysis of correlation between efficiency scores and plant profitability (profitability is measured by EBIT/value of production at retail prices). On the other hand, it estimates regression models (by means of OLS and 2SLS estimators) explaining efficiency and profitability as function of the JIT concentration index and the total number of productivity measures. Measures of productivity are TFP, LP, ROI, Quality of output, Inventory (as total number of productivity measures associated with inventory control), whilst performance outcomes are measured by efficiency and profitability.

The main findings of this study are that productivity measurement mediates the relationship between performance outcomes and intensity of JIT management practices. Specifically, both JIT and non-JIT plants that use a broader range of productivity measures are more efficient and profitable than other plants. Also, plants that employ industry-driven productivity measures are more profitable and efficient than plants that employ idiosyncratic productivity measures, especially if the former are more JIT-intensive than the latter. Furthermore, plants that employ quality productivity measures are less efficient and less profitable than those that do not, especially if they use more intensive JIT practices. This study also finds that, despite the fact that plant profitability and efficiency are highly correlated, JIT-intensive plants are more profitable but less efficient than plants that are not JIT-intensive, after controlling for productivity measures, plant size, and buffer stock. This result suggests that despite wasting resources, JIT-intensive plants are still able to generate relatively higher profits than plants that are not JIT-intensive.

Overall, it seems that there is consensus in the literature about a positive impact of an individual management practice in isolation on productivity.

It is also worth to notice, across the extant literature, that the most commonly used approach to the analysis of the relationship between operational management practices and productivity/performance has been to examine the impact of an individual management practice in isolation on productivity. However, recent theoretical and empirical research suggests that this approach may be misleading since firms often adopt clusters of management practices rather than individual practices in isolation (Ichniowski et al., 1995; Huselid, 1995; Patterson et al., 2004; among others). This is because the presence of complementarities among innovations is such that when an innovation is adopted in isolation it might not necessarily yield positive gains. However, when innovations are jointly adopted they can significantly improve productivity, increase quality and often result in better firm performances than more traditional systems (see for example Ichniowski et al. (1997) and Ruigrok et al. (1999) for applications to human resource management practices or Stoneman
(2004) and Battisti et al. (2005) for theoretical models). In other words, the benefits from the joint adoption of clusters of complementary innovations can be higher than the sum of the individual effects.

Sale and Imman (2003) describe the use of a comprehensive set of criteria to examine empirically changes in business unit performance, over 3 years, as reported by firms adopting JIT and Theory of Constraints (TOC). They compared the performance and the change in performance of companies adopting TOC, those adopting the JIT, those adopting both JIT and TOC, and those reporting to have adopted neither (traditional manufacturing). The methodology used is a Variance analysis to test for performance differences across firms using JIT, those using traditional manufacturing, those using TOC, and those using JIT/TOC. The performance measure reflects 13 potential performance criteria weighted by managers’ importance score, namely: sales level, sales, growth rate, market share, operating profits, profits to sales ratio, cash flow from operations, return on investment, new product development, market development, R&D activities, cost reduction programs, personnel development, political public affairs. Results indicate that the greatest performance and improvement in performance accrued to adopters of TOC. JIT did not have superior performance or superior change in performance when compared with traditional manufacturing. Change in performance for firms, using JIT and TOC, is negative, though not to a significant degree (except when compared with TOC).

Patterson et al. (2004) analyse the impact of a cluster of management practices upon performance by taking into account the possible complementarities between operational and human resource management practices (there is also in the literature some paper, see for example Gale et al., (2002) that explores the relationship between different types of management practices, but not the impact of those practices on performance/productivity). Thus, by distinguishing between Integrated Manufacturing (operational) Practices such as AMT, TQM, JIT inventory control – and Empowerment (human resource management) Practices such as Job enrichment and Skill enhancement – this study analyses three key assumptions:

- Whether operational practices affect human resource management practices.
- Whether operational practices and human resource management practices enhance the company performance.
- Whether there is interaction between operational and human resource management practices.
- In particular, by using a Multiple Regression Analysis, they estimate:
  - The impact of manufacturing practices on empowerment practices.
  - The impact of integrated manufacturing and empowerment practices on subsequent profits and productivity.
  - The impact of integrated manufacturing and empowerment practices only on subsequent productivity.

Common to many studies, the analysis does not use a robust measure of management practices. Criteria of measurement used are essentially qualitative and subjective. Typically, managers are asked to assess, on a certain scale, the extent of use of those practices. This is a major problem as subjective measures are not comparable across firms or even within firms over time. Nevertheless, the measure of productivity/profitability seems to be more robust. Two indices of performance are used: labour productivity (value of net sales per employee) and profits (net financial value of sales deflated by the producer price index).
The empirical results seem to challenge the common view that management practices may affect firm performance/productivity. They show that there is no relationship between integrated manufacturing and empowerment practices and the study did not find any evidence in support of a relationship between the impact of management practices – as integrated manufacturing practices – upon firm performance. These findings raise important questions, as for example: is this result due to the subjectivity of the measurement used? Is this because most companies in their sample have adopted just only a subset of management practices? Or, is it related to the presence of financial constraints, given that most of companies are relatively small? Certainly, whatever will be the correct explanation, it rests the fact that this result questions the findings of the most part of literature and casts doubts on the ability of management practices to positively affect the firm performance.

Even in the case of TQM, a strand of empirical literature – the literature on Lean production whose TQM is a core component – reports that TQM not or only partly explains higher firm’s performance. Lean production principles originate from research into Japanese manufacturing (Womack et al., 1990). These basic principles are: team-based work organisation, active problem solving, high-commitment HRM policies, lean factory practice, (supply chain partnering) tightly-integrated material flows, active information exchange, joint cost reduction, and shared destiny relations.

Oliver et al. (1996), for example, analyse data collected from two international studies involving auto-component manufacturing companies in a total of eight countries: France, Germany, Italy, Japan, Mexico, Spain, the UK and the USA. The questionnaire applied is designed specifically to facilitate the profiling of management practices to determine the extent of use of lean manufacturing practices. Using this single source of information, the study presents evidence that lean production principles partly explain high performance.

Similarly, Lewis (2000), by means of a longitudinal study on lean production applied to the UK, France and Belgium, shows that lean production does not automatically result in improved financial performance. Indeed, being ‘lean’ can restrict the firm’s ability to achieve long-term flexibility.

In another study, TQM exerts little or no observable effect on increasing productivity over the short time it was in place (Kleiner et al., 2002). It in fact reduces labour productivity and increases labour costs, although a positive effect starts to be observed during its second year. It is reasonable to expect that a time lag of some duration is required for a change in management practices to exert an impact, however this study offers initial insights that management under pressure for results are perhaps unable to commit to the achievement of long-term results if the short-term costs are too great.

Black and Lynch (1996), by using a Cobb-Douglas production function and an OLS estimator, estimate, with cross-sectional data, the impact of a cluster of management practices – Human capital and TQM – on productivity. Results of this analysis partially support the finding of some studies that management practices have not impact on productivity. Indeed, even if TQM is found not to have an impact on productivity, human capital management practices affect the productivity of the manufacturing establishment under scrutiny. One of the explanations for these results is that estimated parameters may be not significant because the estimation method does not control for the timing of the introduction of the management practices. In fact, if the firm has only just introduced the management practices, we should expect to see a delay in their impact on productivity, just as the introduction of new physical
capital. Another explanation is that crude measures of the incidence of TQM on productivity do not capture how these programs have actually been implemented. Perhaps, what it is most important is not the introduction of the management practice (TQM) but rather how it has been introduced, when it is introduced and how it has been implemented (this issue has been examined by for example (Ichniowski et al., 2003; Leseure et al., 2004).

3.4 HRM Practices

A strand of literature argues that investment in HRM practices can raise and sustain a high level of firm performance. HRM practices can represent a significant source of competitive advantage because they are the means by which firms to locate, develop and retain rare, non-imitable, and non-substitutable human capital (Barney, 1991; Barney, 2001).

Huselid’s infamous paper (1995) presented empirical evidence in support of a positive, synergistic effect of complementary HRM practices on productivity and corporate performance. This paper is still widely referred to, indicating a lack of subsequent comparable work in the literature; empirical studies that link HRM practices and firm level performance outcomes are sparse (Strategic HR Review 2004). The research design used by these limited number of studies tends to employ overly simplistic analyses. Many authors simply regress a general performance index (for example, value added or sales per employee, or Tobin’s Q) on a set of variables capturing the extent of adoption of HRM practices. Even if HRM practices do raise performance, the statistical significance can be weak.

Ichniowski et al. (1995) formed a statistical distribution of HRM practices to show that some practices are adopted only in presence of some others (i.e. as clusters), and some clusters display a more significant productivity advantage than others. The econometric analysis of this paper is relatively robust as it is based on panel data rather than on cross-sectional data. Building on this finding, Ichniowski et al. (1997) analysed the impact of different clusters of management practices on productivity, to estimate the impact of a single HRM practice on productivity. Empirical results demonstrated that manufacturing lines using a set of HRM practices are associated with a higher level of productivity than lines employing a single HRM practice.

Koch and MacGrath (1996) investigated the impact of a set of HRM practices on labour productivity, to find that investments in HR planning and in hiring practices are positively associated with labour productivity. Results suggest that firms which systematically train and develop their workers are more likely to enjoy the rewards of a more productive workforce than those that do not, although this is not framed to take account of the bigger picture. For example, Cappelli and Neumark (2001) provided some indication that empowering work practices are related to greater productivity. The authors presented partial evidence of such a relationship, however, since the work practices raise labour costs per employee (in this case employee compensation), it is unclear whether such practices are beneficial to the firm overall. Another study, this time of small Belgian companies, revealed a similar situation. Sels et al. (2006) demonstrated a strong and positive relationship between HRM intensity and productivity, controlling for past performance and using one-year lagged financial performance indicators (although the measures were recorded contemporaneously). This beneficial effect was greatly outweighed by the cost increases associated with higher HRM intensity. Nevertheless, HRM intensity was directly related to profitability, and the authors understand this in terms of the minimisation of unmeasured operational issues.
A cross-sectional, single-respondent empirical study of 52 Japanese multinational corporation subsidiaries in the USA and Russia demonstrated that employee skills, attitudes and behaviours play a mediating role between HR systems and firm outcomes (Park et al., 2003). Results suggest that synergistic ‘bundles’ or systems of HR practices positively influence the performance of the types of Japanese subsidiaries concerned. This can be explained in one of two ways: either HRM practices exert an influence regardless of firm location, or Japanese organisations always implement very similar ‘best practices’. Indeed, other empirical evidence suggests that the potential causational path from HRM practices to productivity is more complicated than once thought. Another study, multi-respondent and quasi-longitudinal in design involving Indian software companies presented empirical evidence demonstrating no direct causal relationship between the HRM practices in question and organisational financial performance, although some HRM practices were directly related to operational performance parameters (Paul and Anantharaman, 2003). Instead, it was found that every single HRM practice measured indirectly influenced the organisation’s operational and financial performance. The indirect effect is very important, because few studies employ a research design where intervening variables are measured, but beware that the sample size was too small to apply all of the desired statistical analyses (i.e. maximum likelihood model) and no controls were added. The findings are nevertheless thought-provoking and infer that simply focussing on a direct linkage between HRM and performance may not reveal the operational mechanism through which an effect is exerted.

In support of this type of approach, Michie and Sheehan (2005) analysed original data from a mixed sample of 362 manufacturing and service sector companies. The empirical findings demonstrate positive relationships between HR policies and practices and objective financial performance, mediated by business strategy type (business strategies were classified as cost leadership, innovation-focussed or quality-focussed). Additionally, the use of external flexible labour was associated with lower HR effectiveness. The implications are very pragmatic, and although this survey is only cross-sectional it could be inferred that there exists a two-way causational relationship between the HR policies and practices and financial performance.

Chan et al. (2004) investigated the effects of what they called ‘high performance’ HR practices (some researchers apply loaded terms such as High Performance Human Resource Practices (HPHRP), however there is no scientific basis for this distinction (Wall and Wood 2005). A total of 82 multi-industry firms in Hong Kong were surveyed, and results do not support a relationship between HPHRP and company performance. The authors concluded that the relationship between HRM practices and firm performance may be more complex than originally considered, and suggested that cultural factors may be at work. Apparently companies in Hong Kong tend to focus on short-term results, and HR practices are not as well-established as in other more developed countries. Indeed, this offers an explanation why this Hong Kong study was not very successful. However, significant positive correlations between four out of five organisational culture traits (namely: involvement, policy consistency, adaptability, and mission) and organisational performance are presented, suggesting that the role of organisational culture may be underestimated.

Comparing the productivity of Japanese and US production line workers, empirical evidence indicates that US manufacturers who had adopted a full system of innovative HRM practices patterned after the successful Japanese system achieved levels of productivity and quality equal to the Japanese manufacturers’ performance (Ichniowski and Shaw, 1999). This suggests that the Japanese plants’ average 5% higher productivity cannot be attributed to
cultural differences; instead this is related to the utilisation of more effective HRM practices. Estimating the coefficients of a fixed effects model, Ichnioskwi and Shaw (1999) presented initial evidence from a subset to suggest that a firm transitioning its HRM system to involve more innovative practices will improve productivity. Methodologically this longitudinal study is generally sound. However, counter-intuitively, the approach adopted meant that any time lag for a management practice to impact upon productivity was ignored. For example, if a practice was adopted at the very outset of the study it would have been immediately recorded as present and productivity assessed concurrently.

Bloom and Van Reenen (2006) collected data on 732 manufacturing firms in the UK, France, Germany, and the USA for the period 1992-2004. Data collection involved the application of a novel measurement tool, offering a sophisticated way of assessing and combining ratings of management practices at a grassroots level. Robust estimation techniques were applied, specifically OLS, IV, WG and GMM estimators. The resulting measure of managerial best practice significantly related to sales growth, survival, Tobin’s Q, profitability and productivity. The authors investigated why so many companies survive with relatively inferior management practices, and why this pattern varies so much across the USA and Europe. Findings suggest these phenomena can be explained in terms of low product market competition and eldest sons inheriting control of the family firm (‘primo geniture’). Both of these factors are much more prevalent in the European countries surveyed than the USA, and accounted for around half of the badly managed firms and a similar amount of the inter-continental discrepancy in management performance. The authors also uncovered a large variation of management practices even across firms within each country, especially for the UK.

The authors Bloom and Van Reenen (2006) mitigated the problems associated with a single-respondent design by drawing upon data from three independent sources for each company. Medium-sized firms with 50 to 10,000 employees were targeted to ensure that each company has publicly available financial data (i.e. not too small), each firm uses relatively homogenous management practices and interviewers were unaware of how each company was performing (i.e. not too big). The methodology of this study is commendable, and many different variables are controlled for. However, the universal conceptualisation of particular practices as ‘good’ or ‘bad’ provides only a proxy of management practices and does not allow for the incorporation of context-specific practices that may be more important to other sectors, such as the service sector.

Birdi et al. (2006) investigated the relationships over time between the introduction of seven management practices (JIT, TQM, advanced manufacturing technology, supply-chain management, empowerment, learning culture and teamwork) and audited company performance. Data is presented for a total of 308 companies over a period of 22 years. Results demonstrate a universally positive effect of empowerment on performance, whereas the impact of learning culture appeared to be context-specific. Importantly, the impact of the other five practices varied, indicating that the introduction of a particular management practice can have no or even a negative impact on performance. Statistical relationships between variables were largely incompatible with contemporary theories; a significant finding given that it is highly unlikely these propositions have been previously tested on such a grand scale.

Despite the lack of an individual effect of teamwork, this practice positively interacted with every single other management practice to predict company performance (Birdi et al., 2006),
highlighting how vital teamwork is to a productive organisation. Given the single respondent design, the authors conducted a consistency check and yielded a high consistency rate (84%). It is difficult to criticise this study due to the exceptionally extensive data set and explicit methodology, although the extent of implementation of each practice is not ascertained and it is unclear whether the cessation of practices is incorporated in the analyses. The authors argue that it is likely that only effective practices are institutionalised by an organisation and consequently reported as in use; however this relies upon the assumption effective feedback mechanisms exist to provide accurate information to the organisation’s decision-makers.

High-involvement work practices represent another important set of HMR practices. Employees of a high-involvement organisation take greater responsibility for its success. In practice, this involves human resource practices to develop and support a self-managing and self-programming workforce (Lawler, 1992). Guthrie (2001) received responses from 190 New Zealand companies with at least 100 employees, and empirically demonstrated a positive relationship between the application of high-involvement work practices and productivity. However, an interaction was observed with employee turnover: when productivity was high, employee turnover was linked to decreased productivity; and when productivity low, turnover was associated with increased productivity. Indeed, employee retention is critical when financial investments in work practices are relatively high, and this finding infers that employers may benefit from utilising complementary management practices (such as enhancing retention of good performers) alongside high-involvement systems. This study is dependent on entirely self-reported data from single respondents, and although the practices measurement refers to the previous year all data was collected at the same point in time. It is important to remember that this simultaneity problem does not allow us to eliminate the possibility that an organisation with lower turnover and higher productivity is better positioned to invest in high-involvement work practices.

Bryson et. al. (2005) investigated WERS98 data (this data is collected from a nationally representative sample of organisations using a preferable technique of multi-respondent sampling across organisations (one senior management, one worker/employee representative, and up to 25 employees) for the private sector only to test hypotheses regarding work organisation, trade union representation, and workplace performance. Findings demonstrated a positive effect of HIM practices on labour productivity; however this effect was barely document-able within non-unionised workplaces. Descriptive evidence suggests this effect is attributable to concessionary wage bargaining on the part of unions. Within the union sector, key differences between the characteristics of HIM and more traditional methods of management were identified, suggesting that if a more traditional firm attempted to implement more HIM practices this may not be congruent with the broader organisational context. All in all these research findings raise concerns about the universal applicability of HIM as a method of improving workplace performance. However, this study is only cross-sectional in nature, and any inferences about implementing improvements in productivity can only be considered speculative. Organisations are categorised into groups depending on the number of defined practices adopted, however this presumes every organisational requires the same quantity of practices, neglecting the context-specificity of HRM.

A study within the health service presented empirical evidence from a cross-sectional survey of 112,360 employees. High Involvement Work Systems (HIWS) correlated with lower patient services costs, and therefore improved financial performance (Scotti et al., 2003). Part of the statistical relationship between HIWS and a reduction in costs was mediated by employee satisfaction. In turn, satisfaction was related to a reduced intention to leave, and
fewer instances of other costly organisational outcomes. These results tentatively infer that increasing the use of HIWS reduces organisational costs, although we cannot firmly conclude this from another cross-sectional study.

In a correlational study of 52 manufacturing organisations, effective employee involvement practices were directly associated with employee satisfaction, quality improvement and productivity enhancement (Pun et al., 2001). It is impossible to separate out the causational influences of these variables upon one another; it is apparent there may be some over-arching factor at play such as application of modern management practices, although this is also speculative.

3.5 Miscellaneous Management Practices

3.5.1 Outsourcing

By hiring the service of factors of production from countries that enjoy a comparative advantage (say, a relative abundance of a given input), firms can decrease costs and hence increase productivity. For example if hourly wages in Mexico are lower than in the US, US firms will find it profitable to satisfy part of their labour demand with labourers in Mexico (outsourcing).

The empirical evidence reviewed supports the idea of a positive relationship between outsourcing and productivity. (Egger et al., 2001) uses manufacturing data from Austria to test the productivity impact of outsourcing to Eastern European countries. Their panel data estimation suggests that TFP in Austria increased as a consequence of Eastern European outsourcing occurring between 1990 and 1998. Using an instrumental variable econometric approach for more than 3,000 UK manufacturing firms, Girma and Gorg (2004) find that outsourcing is positively related to labour and total factor productivity.

An excellent literature review on the measurement problems behind the relationship between outsourcing and productivity is found in Heshmati (2003). The author discusses the recent studies in this subject including evidence for the manufacturing as well as the service sector. The great bulk of evidence shows a positive relationship between outsourcing and TFP. Heshmati (2003) stresses the measurement problems faced while estimating TFP and subsequently regressing it against a measure of outsourcing. The author suggests:
- The analysis should be performed at the micro level (firm level analysis)
- Panel data estimation techniques ought to be used
- Identify and estimate the impact of changes in organization and production structure on performance

Control for specific attributes of inputs, outputs, production techniques and other firm-level characteristics (fixed effects estimation)

3.5.2 Business Process Reengineering

Business Process Reengineering (BPR) is a method of reshaping business processes with the goal of enhancing performance. Hammer (1990, p.104) advocates its application, stating that our businesses need to, ‘use the power of modern information technology to radically redesign our business processes in order to achieve dramatic improvements in their performance.’
Rotab Khan (2000) devised BPR for air cargo handling and inferred from calculations that cycle times would be shortened, work efficiencies would be improved, and costs reduced. The author states, ‘BPR has proved to be a modern innovative useful management technique to achieve dramatic improvement in operational efficiencies for quality services of an airline’s cargo handling process,’ (Rotab Khan 2000, p.108) which is surprising considering the proposals had not been implemented at the time of publication. An interesting question would be to ask why; either the company concerned is not convinced of the potential gains, or it is unable to implement the recommendations for some reason. All in all there appears to be a rational basis for the ability of BPR to increase the productivity of air cargo handling, however the in terms of implementation both potential financial and human (for example, one consequence could be a decrease in worker motivation if the scheme is seen as a threat to jobs) costs plus the choice of implementation method have been ignored.

3.6 Conclusions

In this survey we have focussed the attention on HRM and Operational Management Practices – such as Just-in-Time, Total Quality Management and Information and Communication Technology - and productivity/performance. We have observed, on the one hand, how these management practices have been measured and, on the other hand, how the impact of these practices on firms’ productivity/performance has been estimated.

We have also highlighted how productivity, performance and profitability have been measured. The results on the link between management practices and productivity show a number of interesting facts:

- The first fact is that there is consensus about a positive effect of individual management practices (JIT/TQM) in isolation on productivity (or performance). However, when management practices are jointly adopted, there is no consensus on a positive effect of management practices on productivity.
- Second, regarding to the impact of ICT on productivity some studies at industry-level have found a significant either positive or negative contribution of ICT on productivity, while, other studies, using firm-level data, have found positive productivity effect or no productivity effect at all.
- Third, although the econometric methodology appears to be robust and quite sophisticated, a wide range of definitions of management practices and performance/productivity have been used this makes results not robust to comparisons over time and across studies.
- Fourth, most of the studies are carried out at industry level.

These results have some important implications. First, we believe that the lack of universal consensus on the sign of the effect of the adoption of complementary management practices might be driven either by measurement issues or by the level of analysis. In particular it might depend both on how the management practice is defined – given that there is no unique definition of the management practice – and on how we measure management practices and performance. Often the analysis does not rely on a robust measure of management practice/s or productivity. The criteria used often are often qualitative and subjective. Typically, managers are asked to assess, on a certain scale, the extent of use of those practices or the change in productivity over the last year. This is a weakness of this methodology because the main problems arising from using subjective measures is that they are not comparable across firms or even within firms over time. Also, we suspect the lack of consensus is a biased result of the level of analysis. The vast majority of the studies are firm level and only a few studies
have investigated the link between management practices and productivity at plant or establishment level.

The second implication concerns the need of longitudinal data that investigates the firm innovative activity in sectors different from manufacturing, and use levels different from firm or industry level.

For this and other reasons, we strongly believe that there is need for further research. In particular for a multi level approach from the lowest possible level of aggregation up to the firm level of analysis in order to assess the impact of management practices upon the productivity of UK firms.

The prevalence of correlational studies indicates that many researchers are at an early, exploratory stage of trying to understand the mechanics behind how management practices may influence productivity. This type of research design does not facilitate the inference of causality, and is extremely limited in the way it can convey the complexities of relationships between people and processes. Cross-sectional research designs test simultaneous effects, i.e. two-way causal relationship between two variables. A fair number of studies are also limited by small sample size, reducing external validity. Indeed, there are serious concerns about the methodological limitations of research into a link between management practices and productivity, for a thorough review, see Wall and Wood (2005).

Some studies have adopted longitudinal designs with varying success. Indeed, it is more reasonable to conclude that there needs to be some kind of time lag between initial implementation, employee consultation, or union negotiation and the management practices demonstrating some kind of impact on organisational outcomes. It is important to mention here the potential reverse causality of management practices (Savery and Luks, 2004).

The majority of research reviewed has relied upon data collected from single respondents, increasing the chances of common method variance. Undoubtedly, there is an inherent trade-off between reducing common method variance associated with single-respondent designs and ensuring a large enough sample size and sufficiently high response rate to draw generalisable conclusions. It is important to balance the needs of good science with more pragmatic concerns, and appropriate statistical tests can be applied to test for bias prior to subsequent analyses, for example see (Birdi et al., 2006).

Many studies have also relied entirely upon perceptual measures that may incorporate measurement error. However, Wall et al. (2004) empirically demonstrated that perceptual measures of company performance are no less valid or reliable than objective measures. Indeed, there is an argument against using company accounts: accounting conventions and other sources of error may pervade this assumed objective data. It is possible that purely financial performance measures fail to account for the broader organisational picture, therefore the inclusion of non-financial performance criteria such as customer satisfaction, productivity, and quality provide may provide more amenable outcomes. To the contrary, if the bottom line contribution of management practices cannot be demonstrated then their implementation remains highly questionable. A small number of key studies have demonstrated promising linkages between management practices and financial performance (Michie and Sheehan, 2005; Paul and Anantharaman, 2003).
Management practices are multi-dimensional constructs that generally do not demonstrate a
straight-forward relationship with productivity variables. Empirical evidence suggests that
effective management practices need to be context specific, as productivity indices need to
reflect a particular organisation’s activities. Consequently it is tricky to ascertain whether the
finding of a relationship, or no relationship, is a fair conclusion. Some researchers have risen
to the challenge and adopted more sophisticated methods of operationalisation, for example
Bloom and Van Reenen (2006) and analysis that offer greater scope for unravelling the
complex interrelated and mediationary relationships at play. For example, one study
uncovered a curvilinear relationship between management practices and performance (Maes
et al., 2005), indicating that beyond a certain amount or intensity management practices
actually diminish performance. Correlational research to ascertain relationships between
other workplace constructs and productivity may help inform future research into mediation,
such as Geralis and Terziovski (2003) or Silvestro (2002).

There is a fair amount of support for a contingency approach; however it is unclear what the
common factors to consider are, see Birdi et al. (2006). Applying context-specific measures
creates variability between research findings and renders them directly incomparable. For
example, it is apparent there are contrasting definitions of lean production techniques, and
these difficulties in achieving consensus makes it likely that each firm follows a ‘unique lean
production trajectory’ (Lewis, 2000; p. 975). Whereas on the other hand, TQM practices tend
to be involve a similar set of practices within whichever organisation they are implemented
within. Indeed, there remains scope for the future investigation of degrees of internal,
organisational and strategic fit (Wall and Wood, 2005).

Taken as a whole, the research findings are equivocal. Some studies have found a positive
relationship between management practices and productivity, some negative and some no
association whatsoever. It is apparent that applying terminology such as ‘high performance
work systems’ is very presumptuous, and more attention needs to be paid to contextualised
working practices and the mechanisms through which these practices may impact upon
organisational performance.

4. Operational Research Models of Management Practices

4.1 Introduction

So far the different Operational and HR management practices and their indicators have been
identified in the literature and their impact has been reviewed. Furthermore, ways to measure
them and typical performance measures to assess firm productivity have been discussed.
Measurement difficulties have been outlined. It has been concluded that management
practices are multidimensional constructs that generally do not demonstrate a straightforward
relationship with productivity variables. Also, empirical evidence suggests that effective
management practices need to be context specific, as productivity indices need to reflect a
particular organisation’s activities.

In order to support managers when they make decisions about implementing management
practices in their company, Operational Research (OR) can be employed to provide decision
support in form of decision support tools. OR is a scientific approach in managing
organisations (Hillier and Lieberman, 2005). In general OR is applied to problems
concerning the conduct and co-ordination of the operations within an organisation. It has
been used in industry since the early 1950’s in response to the growth in size and complexity
of organisations.
An OR study usually involves the development of a scientific model that attempts to abstract the essence of the real problem. When investigating the behaviour of complex systems the choice of an appropriate modelling technique is important.

As outlined in previous sections, we are interested in understanding and predicting the impact of different management practices on retail store productivity. In order to achieve this goal we have planned to follow a step-by-step strategy. In a first step we focus on modelling individual departments of a department store. Later we will expand our horizon and model different department stores. Ultimately, we aim to design a model that allows a cross-country comparison (UK / USA).

With regards to the output of the study we are primarily interested in system performance (e.g. sales, staff utilisation, profitability) but also in capturing emergent phenomena. Emergence occurs when interactions among objects at one level give rise to different types of objects at another level. More precisely, a phenomenon is emergent if it requires new categories to describe it which are not required to describe the behaviour of the underlying components (Gilbert and Troitzsch, 2005). Capturing emergent phenomena is necessary to not just measure, but understand the impact of the management practices.

There are three questions we have used to structure this part of the literature review:
1. Which areas are worth investigating in order to find a suitable modelling approach?
2. Do any models currently exist, that model the impact of management practices on productivity in the retail sector?
3. What modelling approaches are commonly used within related areas?

We have included 196 documents in order to find answers to the questions raised. From these we have identified 28 key documents which we cite throughout the text. The next section includes a discussion of these questions while the last section presents the conclusions from the review and gives some ideas regarding our future plans.

**Figure 1. Review areas, modelling approaches, and application types**
4.2 Review of Relevant Literature

4.2.1 Question 1: Which areas are worth investigating to find a suitable modelling approach?

With our focus on understanding and predicting the impact of different management practices on retail store productivity, we have searched the literature for applications and suitable modelling methods. An overview of the areas and techniques that we have considered for the review is presented in Figure 2.

4.2.2 Question 2: Do any models exist, that model the impact of management practices on productivity in the retail sector?

After defining the areas of interest we extensively reviewed material available on the subject, with particular emphasis on those publications that investigate the link between management practices and productivity in the retail sector. We found a limited number of papers that investigate management practices in retail at firm level. The majority of these papers focus on marketing practices, (e.g. Cao, 1999; Keh et al., 2006). One noteworthy exception is Berman and Larson (2004), who investigate the efficiency of cross trained workers in stores. By far the most frequently used modelling technique in all these papers is agent-based modelling which will be discussed in more detail later. It seems to be the natural way of system representation for these purposes.

Most papers that investigate retail productivity focus primarily on consumer behaviour and efficiency evaluation with less emphasis on retail management practices. These papers can be further segmented into those investigating high street retailing and those focusing on online retailing, i.e. physical and electronic distribution channels. An advantage of looking at online retailing is the availability of data, due to every click being recorded. On the other hand, data in many high street stores is available from loyalty cards, credit cards, sales slips and customer surveys. An interesting contribution is made by Nicholson et al. (2002), who compare different marketing strategies for multi channel (physical and electronic) retailing. They conclude that there are big differences within the consumer decision making process in terms of the different channels.

Regarding the investigation of cross-country differences in relation to the application and impact of retail management practices on productivity, we found no papers that describe attempts to model this issue.

In terms of commercial software, an example was found which simulates the relationship between certain management practices and productivity. ShopSim (Savannah Simulations, 2006) is a decision support tool for retail and shopping centre management. It evaluates the shop mix attractiveness and pedestrian friendly design of a shopping centre. The software uses an agent based approach, where behaviour of agents depends on poll data. It is a good example for the form the output of our OR study could take. Only, in our case the tool would operate on a department level rather than on a shopping centre level and would investigate different kinds of management practices, rather than a shopping centre layout. Furthermore, the input data would come from management and staff surveys rather than from customer polls.

To summarise, we can say that to date only limited work has been conducted in developing models that allow investigating the impact of management practices on productivity in the
retail sector. Therefore, in the next section we broaden our view and review different modelling techniques that are commonly used in the areas of interest identified earlier in response to Question 1. We also include other sectors in our review (e.g. manufacturing sector and other service sectors) to investigate the methods used there to understand and predict system behaviour.

4.2.3 Question 3: What modelling approaches are commonly used within related areas?

The relevant modelling approaches can be divided into three different categories: analytical approaches, heuristic approaches, and simulation. In this section we first compare these approaches in a general sense and then continue to discuss the most important methods individually, describing what they are, where they are normally used, and how they fit in with our OR study. Table 1 presents a summary of the number of relevant papers found in each category for different sectors.

Table 10. Relevant papers found, sorted by modelling approach and application sector

<table>
<thead>
<tr>
<th>Approach</th>
<th>Sector</th>
<th>Manufacturing</th>
<th>Retail</th>
<th>Catering</th>
<th>Banking</th>
<th>Other</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Envelopment Analysis</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Regression &amp; Cluster Analysis</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Bayesian Nets</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Petri Nets</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Queuing Networks</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Structural Equation Modelling</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Probabilistic Approach</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Stochastic Frontier Functions</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>14</td>
<td>20</td>
<td>1</td>
<td>29</td>
<td>7</td>
<td>5</td>
<td>31</td>
</tr>
</tbody>
</table>

From the occurrences within the literature Table 10 indicates that Agent-Based Modelling and Simulation (ABMS) and Data Envelopment Analysis (DEA) are the most frequently used modelling techniques within our area of interest. While most of the ABMS papers studied relate to the retail sector most of the DEA papers are from the manufacturing sector.

In many cases it was found that a combination of modelling techniques was used within a single model. Where this is the case the individual modelling techniques have been listed in the table. Common combinations are simulation / analytical for comparing efficiency of different non-existing scenarios, (e.g. Greasley, 2005), and simulation / analytical or simulation / heuristic where analytical or heuristic models are used to represent the behaviour of the entities within the simulation model (e.g. Yildizoglu, 2001; Schwaiger and Stahmer, 2003).

4.2.4 Analytical Modelling Approaches

Once data has been collected it is common in economics and social science to use analytical analysis tools to quantify causal relationships between different factors. Often some form of regression analysis is used to investigate the correlation between independent and dependent variables. A good example of this type of analysis can be found in Clegg et al. (2002) who
investigate the use and effectiveness of modern manufacturing practices. Survey data is analysed using parametric and non-parametric analytical techniques, as appropriate to the nature of the response scales and the distributions of scores obtained. The results of these analyses are then often used as a data source for heuristic and simulation models.

A number of papers were found that use different analytical modelling approaches to investigate very specific individual hypotheses in relation to management practices and consumer behaviour. A noteworthy paper is by Patel and Schlijper (2004) who use a multitude of different analytical and other modelling approaches to test specific hypotheses of consumer behaviour.

Another area where analytical models have often been used is the assessment of relative efficiency of comparable alternatives. The most frequently used techniques described in the literature are DEA and Multi Level Modelling (MLM). DEA was originally used for assessing the efficiency of the public sector, but has since evolved and is in use in other sectors, often as an analysis instrument for the output of simulation experiments. The objective of DEA is to assess the relative efficiency of a number of decision making units using a variety of input and output data. It is a benchmarking technique which has gained increasing popularity during the last few years. An interesting application of DEA can be found in Soteriou and Stavrinides (1997) where customer service quality, operating efficiency and profitability between different bank branches is compared. MLM is a form of hierarchical regression analysis, designed to handle hierarchical and clustered data. It enables one to contextualise the raw outcomes taking into account the different levels of aggregation in the data. Most examples reviewed were used to assess educational services like schools (e.g. Goldstein, 2003) and universities (e.g. Johnes, 2003).

Major differences between DEA and MLM are explained in Thanassoulis et al. (2003). The main difference is that DEA is a non-parametric method where the focus in each instance is the unit being assessed while MLM is a parametric method seeking to establish a functional relationship at a global level across all units being assessed. Another major difference is that DEA is a deterministic method that does not impose any statistical distribution to the random noise while MLM does so and allows such noise to affect the estimates at different hierarchical levels. Finally, MLM predicts average performance for given contextual data while DEA is a boundary method estimating best performance for given contextual data. Comparisons have shown that DEA and MLM performed on individual data provide different measures of efficiency (Johnes, 2003). It appears that the level of analysis (individual or firm) is also important. Use of both methods in parallel for validation can draw attention to potential areas where each method can lead to biased results so that user judgement can be brought to bear in arriving at policy decisions.

4.2.5. Heuristic Approaches

No purely heuristic models were found during the review. This does not come as a surprise as pure heuristic models are more frequently used in system optimisation which is not the focus of our OR study. However, heuristic models are often used in combination with other modelling techniques as mentioned earlier and so in these cases they have been included in Table 1. As such they will not be investigated further in this review.
4.2.6. Simulation Approaches

Simulation introduces the possibility of a new way of thinking about social and economic processes, based on ideas about the emergence of complex behaviour from relatively simple activities (Simon, 1996; Gilbert and Troitzsch, 2005). Implemented specifically as a method of theory development, simulation allows clarification of a theory and investigation of its implications. While analytical models typically aim to explain correlations between variables measured at one single point in time simulation models are concerned with the development of a system over time. Furthermore, analytical models usually work on a much higher level of abstraction than simulation models. For simulation models it is critical to define the right level of abstraction. Csik (2003) states that on the one hand the number of free parameters should be kept on a level as low as possible. On the other hand, too much abstraction and simplification might threaten the homomorphism between reality and the scope of the simulation model.

There are several different approaches to simulation. The choice of the most suitable approach always depends on the issues investigated, the input data available, the level of analysis and the type of answers to be sought.

Discrete Event Simulation (DES) can be described as the modelling of a system as it evolves over time by use of a representation in which the state variables change instantaneously when an event occurs. It is often used in OR and Operations Management for modelling stochastic dynamic systems, for example the design and operation of queuing systems, management of inventory systems and health care applications. It is not well suited to modelling human oriented systems, where the behaviour of individuals has an impact on system state development over time (Siebers, 2006). This is due to the fact that DES does not support the concepts of autonomy and pro-activeness, which are important properties required to represent individuals in an appropriate way. Komashie and Mousavi (2005) present an example where DES can be used to model a human oriented system. This is an exception as it is a scheduling problem where humans are just resources.

System Dynamics (SD) has its roots in systems of difference and differential equations. A target system with its properties and dynamics is described using a system of equations that derive the future state of the target system from its actual state. SD is restricted to the macro level in that it models the target system as an undifferentiated whole. It is frequently used in economics to predict the behaviour of complex non-linear systems, for example the growth of an economy on the macroeconomic level (Gilbert and Troitzsch, 2005). As SD considers only one level and the study on emergence requires at least two levels this simulation technique is not suitable for studying emergent behaviour. An and Jeng (2005) provide an example of the application of SD to Business Process Modelling, in this case for supply chain management.

Micro Simulation (MS) is a modelling technique often used to analyse the effects of financial and social policy interventions. It allows the consideration of two levels, which is necessary as researchers are mostly interested in the effects of policy interventions at the aggregate level. An example is the effect on tax income of changing the tax system as a whole; however this aggregate tax income critically depends on individual effects or more precisely on the income distribution of the households observed at the microeconomic level. A representative sample of a population will easily contain several 1000 households because its sub-samples with respect to all property combinations relevant to the purpose of prediction must be large enough to allow projection (Gilbert and Troitzsch, 2005). An example of a
typical application of this technique is given by Sutherland (2004), who uses MS to investigate the impact of government policy on poverty development in Britain.

Agent Based Modelling and Simulation (ABMS) was identified earlier as a key modelling technique for the kind of research questions to be investigated in our OR project. It is therefore discussed in a bit more detail here.

Agent-based modelling which appeared first in the early 1990s is described by Jeffrey (2003) as a mindset as much as a technology: “it is the perfect way to view things and understand them by the behaviour of their smallest components”. ABMS can be used to study how micro level processes affect macro level outcome. A complex system is represented by a collection of agents that are programmed with simple behaviour rules. Agents can interact with each other and with their environment to produce complex collective behaviour patterns. Agent based modelling is using a bottom-up approach as the system is described from the point of view of its constituent units, as opposed to a top-down approach, where you look at properties at the aggregate level. Macro behaviour is not simulated; it emerges from the micro-decisions of individual agents. The main characteristics of agents are their autonomy, their ability to take flexible action in reaction to their environment and their pro-activeness depending on motivations generated from their internal states. They are designed to mimic the attributes and behaviours of their real-world counterparts. The system’s macro-observable properties emerge as a consequence of these attributes and behaviours and the interactions between them. The simulation output may be potentially used for explanatory, exploratory and predictive purposes (Twommey and Cadman, 2002).

Due to the characteristics of the agents mentioned this modelling approach seems to be more suitable than DES for modelling human oriented systems. ABMS seems to promote a natural form of modelling, as active entities in the original are also interpreted as actors in the model. There is a structural correspondence between the real system and the model representation, which makes them more intuitive and easier to understand than for example a system of differential equations as used in SD. Hood (1998) adds that one of the key strengths of ABMS is that the system as a whole is not constrained to exhibit any particular behaviour as the system properties emerge from the constituent agent interactions, so assumptions of linearity, equilibrium and so on, are not needed.

With regards to disadvantages there is general consensus in the literature that it is difficult to evaluate agent-based models, as the behaviour of the system emerges from the interaction between the individual entities. Carley (1996) argues that for “intellective models” (e.g. models that illustrate the relative impact of basic explanatory mechanisms) validation is somewhat less critical and the more important thing is to maintain a balance between keeping a model simple and maintaining veridicality (veridicality is the extent to which a knowledge structure accurately reflects the information environment it represents). Furthermore, problems often occur through the lack of adequate data. Twommey and Cadman (2002) state that most quantitative research has concentrated on ‘variable and correlation’ models that do not cohere well with process-based simulation that is inherent in agent based models. A final point to mention is the danger that people new to ABMS may expect too much from the models, particularly in regard to predictive ability.

Models used in Agent-based Computational Economics (ACE) are a good example of the application of the agent paradigm with relevance to our OR study. ACE is the computational study of economies modelled as evolving systems of autonomous interacting agents to study
the evolution of decentralized market economies under controlled experimental conditions (Tesfatsion, 2003). An example of such a study can be found in Vriend (1995) where the emergence of self-organized markets in a decentralized economy is investigated using ABMS.

Marketing Science is another area where ABMS is used increasingly to describe, model and predict the behaviour of consumers and their attitudes towards the products of the market. Csik (2003) for example, investigates the impact of micro- and macro-level driving factors that affect consumer behaviour and strategies of firms to influence people’s behaviour. The final example to be mentioned here is crowd behaviour modelling, which seeks to find emergent patterns in crowd movements. Kitazawa and Batty (2004), for example, investigate the retail movements of shoppers in a large shopping centre.

4.3 Conclusions

In this review, we were looking for modelling techniques best suited to understand and predict the impact of different management practices on retail store productivity. Our current focus is on modelling an individual department store.

While deciding about the modelling technique it was important to keep a few general points in mind. These were: the availability of required data, the allowed level of abstraction, the number of system levels needed to capture emergence, and the possibility to evaluate the developed models.

Firstly we investigated whether any models exist for this particular purpose. The review has shown that to date only limited work has been conducted in this field. The papers we found primarily focused on consumer behaviour and efficiency evaluation with less emphasis on retail management practices. We did not find any attempts to model the issue of cross-country differences. One software package was identified which is able to simulate the relationship between certain management practices and retail productivity. Contrary to the modelling tool that we need for our enquiries, the tool is operating on shopping centre level and is investigating a different kind of management practices.

We then broadened our view and reviewed different modelling techniques that are commonly used in other sectors to understand and predict system behaviour. We found that for our area of interest ABMS and DEA were the most frequently used modelling techniques. While ABMS is a simulation method used for understanding system behaviour, DEA is an analytical method used as a benchmarking tool and more and more often also used as an analysis instrument for simulation experiments. In many cases, it was found that a combination of modelling techniques was used within a modelling study. Often DES and DEA were used together to allow different kinds of analyses. Furthermore, analytical or heuristic models were frequently used inside the agents of agent-based models to give them, for example, decision-making or learning capabilities. Those studies that attempted to model management practices were predominately concerned with marketing practices. Usually the models were designed to answer a specific question, focusing on a particular marketing practice, and customers were modelled in great detail, including, for example, psychological factors.

As a result of the literature review, we have decided to start our modelling efforts using ABMS in conjunction with Bayesian Reasoning for the decision making of our agents. Once we have developed the first models, we will decide whether we require a second modelling technique to compare the relative efficiency of the simulation results. In this case, we would use DEA as a benchmarking tool.
5. SUMMARY

5.1 ASSESSING THE GAP
There seems to be a gap between the UK and other countries (Germany, France, UK), but the size, sectors and reasons for the gap are a matter of ongoing debate. It is not advisable to underestimate the bias caused by methodological choices when evaluating the gap or a country’s performance. Different studies draw very different pictures about the current situation. Data shows a picture of a strong UK economy performing well over the last decade with higher GDP per hour growth rates than all its competitors. That the gap still exists is because the countries had different historical starting positions.

5.2 MEASURES ISSUES
The economic literature has used two different definitions of productivity. The first and more straightforward one, called partial factor productivity, captures the amount of output per labourer or hours of work. The second approach for productivity measurement is TFP, better understood as an index number measuring the change in production that is not explained by the change in inputs. Despite all the econometric and measurement problems associate with TFP, this measure of productivity is the preferred one.

There is no consensus in the literature regarding how to measure management practices. The only commonality shared by all studies is that management practices are measured in a multidimensional fashion. Due to the inherently intangible nature of management practices, it is challenging to apply objective forms of measurement. Measures are aggregated to facilitate analysis either at the plant, firm, industry or country level.

One of the major shortcomings of the studies in this area is the lack of sufficiently long longitudinal data at firm (or plant) level. This has, indeed, severely constrained the possibility to test in a robust way the presence of complementarities between two innovations; the direction of their adoption sequence and the lagged impact of the adoption decision upon firm performance. Several studies use cross-sectional analysis. However, the latter may be biased because of endogeneity problems, i.e. the presence of unobserved firm characteristics that are time invariant.

5.3 MANAGEMENT PRACTICES
The link between management practices and productivity show the following facts:
- The first fact is that there is consensus about a positive effect of individual management practices (JIT/TQM) in isolation on productivity (or performance). However, when management practices are jointly adopted, there is no consensus on a positive effect of management practices on productivity.
- Second, regarding to the impact of ICT on productivity some studies at industry-level have found a significant either positive or negative contribution of ICT on productivity, while, other studies, using firm-level data, have found positive productivity effect or no productivity effect at all.
- Third, although the econometric methodology appears to be robust and quite sophisticated, a wide range of definitions of management practices and performance/productivity have been used this makes results not robust to comparisons over time and across studies.
- Fourth, most of the studies are carried out at industry level.
These results have some important implications. First, we believe that the lack of universal consensus on the sign of the effect of the adoption of complementary management practices might be driven either by measurement issues or by the level of analysis.

The second implication concerns the need of longitudinal data that investigates the firm innovative activity in sectors different from manufacturing, and use levels different from firm or industry level. The majority of research reviewed has relied upon data collected from single respondents, increasing the chances of common method variance.

Management practices are multi-dimensional constructs that generally do not demonstrate a straight-forward relationship with productivity variables. Empirical evidence suggests that effective management practices need to be context specific, as productivity indices need to reflect a particular organisation’s activities.

Results of the analysis show that, consistently with the predictions of the theoretical model, a worsening of the firm performance – due to a reduction in productivity or price of output - will lead to the introduction of management innovations both for firms and plants, and that firms with higher debt burden (under financial pressure) are less likely to introduce management innovations.

Taken as a whole, the research findings are equivocal. Some studies have found a positive relationship between management practices and productivity, some negative and some no association whatsoever. It is apparent that applying terminology such as ‘high performance work systems’ is very presumptuous, and more attention needs to be paid to contextualised working practices and the mechanisms through which these practices may impact upon organisational performance.

5.4 OPERATIONAL RESEARCH AND MANAGEMENT PRACTICES

In order to support managers when they make decisions about implementing management practices in their company, Operational Research (OR) can be employed to provide decision support in form of decision support tools. When investigating the behaviour of complex systems the choice of an appropriate modelling technique is very important. Different approaches will suit different purposes and answer different questions.

By far the most frequently used modelling technique in all these papers is agent-based modelling. It seems to be the natural way of system representation for these purposes. Most papers that investigate retail productivity focus primarily on consumer behaviour and efficiency evaluation with less emphasis on retail management practices.

Regarding the investigation of cross-country differences in relation to the application and impact of retail management practices on productivity, we found no papers that describe attempts to model this issue. To summarise, we can say that to date only limited work has been conducted in developing models that allow investigating the impact of management practices on productivity in the retail sector.

Table 10 indicates that Agent-Based Modelling and Simulation (ABMS) and Data Envelopment Analysis (DEA) are the most frequently used modelling techniques within our area of interest. While most of the ABMS papers studied relate to the retail sector most of the DEA papers are from the manufacturing sector.
As a result of the literature review, we have decided to start our modelling efforts using ABMS in conjunction with Bayesian Reasoning for the decision making of our agents. Once we have developed the first models, we will decide whether we require a second modelling technique to compare the relative efficiency of the simulation results. In this case, we would use DEA as a benchmarking tool.

References


ESRC Seminar Series (2004), “Mapping the public policy landscape. The UK productivity gap. What research tells us and what we need to find out”.

Evidence from UK establishment-level data”, *Review of International Economics*.


models, (3rd Ed).


**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2SLS</td>
<td>High Involvement Work Practices</td>
</tr>
<tr>
<td>ABMS</td>
<td>Agent Based Modelling and Simulation</td>
</tr>
<tr>
<td>ACE</td>
<td>Agent-based Computational Economics</td>
</tr>
<tr>
<td>AIM</td>
<td>Advanced Institute of Management Research</td>
</tr>
<tr>
<td>AMT</td>
<td>Advanced Manufacturing Technology</td>
</tr>
<tr>
<td>BPR</td>
<td>Business Process Reengineering</td>
</tr>
<tr>
<td>CES-TL</td>
<td>Data Envelopment Analysis</td>
</tr>
<tr>
<td>DEA</td>
<td>Discrete Event Simulation</td>
</tr>
<tr>
<td>DES</td>
<td>Dynamic Panel Data</td>
</tr>
<tr>
<td>DTI</td>
<td>Department of Trade and Industry</td>
</tr>
<tr>
<td>EBIT</td>
<td>Gross Domestic Product</td>
</tr>
<tr>
<td>EVA</td>
<td>Gross Margin Return on Labour</td>
</tr>
<tr>
<td>GDP</td>
<td>High Involvement Work Systems</td>
</tr>
<tr>
<td>GMOL</td>
<td>High Performance Human Resource Practices</td>
</tr>
<tr>
<td>HIM</td>
<td>Human Resources</td>
</tr>
<tr>
<td>HIWS</td>
<td>Human Resource Management</td>
</tr>
<tr>
<td>HPHRP</td>
<td>Human Resource Management Practices</td>
</tr>
<tr>
<td>ICT</td>
<td>Information and Communication Technology</td>
</tr>
<tr>
<td>IT</td>
<td>Information Technology</td>
</tr>
<tr>
<td>IV</td>
<td>Just In Time</td>
</tr>
<tr>
<td>KPI</td>
<td>Key Performance Indicators</td>
</tr>
<tr>
<td>LP</td>
<td>Multi Level Modelling</td>
</tr>
<tr>
<td>MLM</td>
<td>Micro Simulation</td>
</tr>
<tr>
<td>OR</td>
<td>Operational Research</td>
</tr>
<tr>
<td>ROCE</td>
<td>System Dynamics</td>
</tr>
<tr>
<td>SD</td>
<td>Total Factor Productivity</td>
</tr>
<tr>
<td>TOC</td>
<td>Theory Of Constraints</td>
</tr>
<tr>
<td>TQM</td>
<td>Total Quality Management</td>
</tr>
<tr>
<td>WG</td>
<td>Within Groups</td>
</tr>
</tbody>
</table>
The Unintended and Indirect Effects of Performance Measurement and Regulation on UK Productivity: A Multidisciplinary Overview


Abstract
This paper provides a multifaceted, interdisciplinary literature review of the indirect, unintended and unforeseen consequences of regulation and performance measurement on productivity. First, the paper analyses the literature on regulation and on performance measurement in relation to regulation. This is followed by three case studies that examine regulation across a different empirical dimension: one focuses on a specific type of legislation (environmental), one on a specific industry (engineering) and the final one on a specific regulatory event (the deregulation of air travel).

Contents
1. Introduction (Gerben Bakker and Stavroula Iliopoulou).
2. The unintended and indirect effects of regulation on productivity: Economic background and cases studies (Gerben Bakker and Stavroula Iliopoulou).
3. Performance measurement, regulation and UK productivity: A multidisciplinary overview of unintended and indirect effects (Kim Tan and Rosalind Rae).
5. An exploratory investigation into the impact of legislation and regulation on industrial organisations: Tentative speculations from reviewing ‘The Engineer’ (Naomi Brookes and Kathryn Walsh).
6. Air travel de-regulation and local economies in Europe: Reflections on the unintended consequences for productivity (Vlado Balaz and Allan Williams).

Introduction
Gerben Bakker and Stavroula Iliopoulou

Considerable research has been done on regulation and its effects. Numerous research centres and policy units are exclusively devoted to regulation. The ESRC [2004] report has identified regulation as one of the possible sources of the UK productivity gap. Assuming that the aim of regulation never is to decrease or impede productivity growth per se, this negative effect of regulation on productivity can be seen as indirect, unintended and/or unforeseen per definition. Analyzing existing literature, the present article examines this effect and dissect it into its various constituent parts.

In short, this paper aims to study the literature on the indirect, unintended and unforeseen consequences of regulation for productivity growth, with special attention to the light this can
shed on the UK case. Three important characteristics stand out. First of all, this paper is interdisciplinary: it is written by experts from several different disciplines, and analysis literature from several different disciplines. Second, its contribution tends more to the practical than to the theoretical, as its aim is, by analysing the literature, to arrive at some conclusions that give insight into the UK situation and into potential areas that warrant larger future empirical and theoretical research. The aim of the study is not to build watertight theoretical positions or engage in textual exegesis of definitions, but to arrive at some practical insights. Third, the research is strongly informed by specific case studies as focal points to examine the interaction of indirect, unintended and unforeseen effects.

This research is worthwhile because of three main reasons. First, regulation and performance measurement is of critical importance for understanding productivity performance [Morgan and Engwall 1998]. Hahn and Hird [1991], for example, found that the overall costs of regulation to the US economy in 1988 amounted to 0.8 percent of GDP, while the redistributive effect amounted to 4.4 percent of GDP. The over-all influence of regulation within the economy was far larger, as many costs and benefits for different industries/areas cancelled each other out when net national impacts were measured.

Second, within the debate on the effects of measurement and regulation on productivity, the literatures that exist within different disciplines have hardly been connected, and have hardly ‘talked to each other’ (a rare exception being Noll [1985]). Several literatures exist in disciplines such as anthropology, sociology, psychology, economics, operational management and strategy, economic geography and planning, industrial economics, engineering and organizational behavior.

Third, the project is innovative in the sense that it is multidisciplinary and aims to bridge the gaps between disciplines, in that it is creative and adventurous with the possibility of unexpected outcomes, and in that it is led by five researchers from different UK universities, each of whom usually works in a different discipline.

The rest of this paper is structured as follows. The next section gives an overview of the empirical research done in the broad economics and social science discipline on the indirect effects of regulation, and examines this by country, industry, type of regulation and time period (by G. Bakker and S. Iliopoulou). The subsequent section moves down to the level of the organization and investigates the commonalities and differences between the indirect effects of regulation (by G. Bakker and S. Iliopoulou) and those of performance measurement (by K. Tan and R. Rae) identified in the literature, addressing, among other things, various theories of incentives. These two general sections are followed by three case studies, which were developed, focused and designed so that each one covers the indirect effects of regulation across a different empirical dimension. The first case study is on a specific type of regulation, the second on is on a specific industry, and the third one is on a specific regulatory event: Section 4 examines the literature on the indirect effects of environmental regulation (by S.J. Antony and A. Gurtoo), section 5 deals with the indirect effects of regulation on the engineering industries using an analysis of the trade press and trade magazines (by N. Brookes and K. Walsh), and the final section 6 examines the literature on the indirect effects of airline deregulation on the UK economy (by A. Williams and V. Balaz).
The effects of regulation on productivity: A literature study of indirect and unintended effects by time, place and industry

Gerben Bakker and Stavroula Iliopoulou

Introduction

Among the prevailing views on regulation, one defines it as being fundamentally a politico-economic concept used as a channel for public agencies to exercise sustainable and focused control over activities that affect and are valued by communities. Although regulation has been researched in the UK, following the recent reports on the existence of a UK “productivity gap”, it is alleged that, regulation may still be a possible cause for shortfalls in productivity. Two documents from DTI (1997 and 1998) set out the British government’s views on the reasons for such a productivity gap in the UK. Among the key points highlighted in the documents indicate that in order for the UK to be regarded as one of the strongest economies, the government should focus on improving UK productivity.

This in turn will improve costs, increase profitability and lead to better rewards for human capital inputs. The importance of investigating the nature of a possible relationship between regulation and productivity lies in the fact that productivity is a key indicator of economic health, and regulation should generally aim for improvement or preservation of the prosperity of an economy, and, or society. So, prosperity or depression may result when regulation aims at the four main avenues recognized as drivers of productivity growth: the role of competition; capital investment; skills and human capital; and innovation. Formally, changes in total economic welfare seem the best indicator of changes in the economic health of a society. In the competitive situation, however, high productivity will result in lower prices and higher profits and therefore larger total welfare. Even in imperfectly competitive industries one could argue, using the Schumpeterian argument that in the long run, higher productivity will result in higher total welfare, as innovators are imitated and monopoly profits invested in new innovations.

Consequently, looking at what is known or what has been established by research about the linkage between UK governance arrangements (for example attempts to carve the structure of markets), competition (such as through state control over resources or activities and their prices) and performance (for example through the adoption of new technologies and the elimination of slack in the use of inputs) could provide essential clues about the different ways in which regulation can improve (or obliterate) productivity in the UK.

In an attempt to investigate this relationship, this paper provides an overview of the literature on the effects of regulation on productivity, and particularly it’s unintended, unforeseen and/or indirect consequences. Our main question is whether, first of all, such effects exist and, further, what is their role in shaping productivity. To address this question, we analyze and critically evaluate published literature that identify that regulation has an impact on the economy that can partially (or entirely) be attributed on effects of indirect, unintended and, or unforeseen nature. In other words, our first contribution is to offer the first study (to our knowledge) attempting to collectively assess the nature and scale of such effects apparent from the existing literature, both individually and collectively, by presenting their possible (inter-) relations and their likelihood to occur. This provides a clearer illustration of the
variety of effects and their impact, and makes a step forward in discussing the possible extent that a regulatory system can (if at all) diversify against such effects.

An additional contribution is that we concentrate not only on the success or failure of regulation to address and achieve its prime targets, but also (after researching an extant amount of literature) the identification and presentation of a collection of existing case studies in the literature that illustrate the several forms and magnitudes of such effects born by regulatory change on measures of productivity. We present some 46 case studies. However, the number of case studies may change since this is work in progress. We estimate that some 15 more cases may be added to the study by the time it is completed. that assess the indirect, unintended and, or unforeseen effects and analyze their evolution across time and geographical space, in an attempt to highlight possible gaps in UK research (of specific regulatory interventions that have not been empirically investigated, unexplored periods of time, or under-researched economies) that cry for some or further attention. Therefore, our paper offers a basis and stimulus for further research in understanding the impact of regulation and may result to an increasing awareness around the impact of regulation on productivity, which may benefit both those interested in regulatory planning and those directly affected by both by the impact of any change brought by regulation and its possible (mainly long-term) side-effects.

The remaining of the paper is organized as follows. The section that follows discusses the different effects of regulation and the way these may relate to productivity. Section 3 presents our research strategy. Section 4 offers a descriptive summary and an analytical discussion on the 46 case studies, respectively, and the final section concludes.

**Regulation and its effects**

Over the last 40 years there has been a considerable amount of research, carried out internationally, attempting to investigate the scope and the forces behind regulation and to evaluate the effect of regulation on the interested and other communities and their operations.

Along the lines of this research, the prevailing models highlight either the orthodox economic view where regulation is perceived as a means used by public authority in an attempt to curb abuses arising from market imperfections, or the views of more skeptic academics like Stingler (1971) pointing out that regulation is closely linked to rent extraction since it primarily intended to benefit organizational interests and safeguard their strive for survival and expansion. Under this model, rent-seeking is more efficient for smaller and more concentrated (organizational) interest groups. However, another upsurge of researchers supports that regulation, apart from the two extremes described above, can rather lie somewhere in their middle, implying that, while the regulated industry (or its’ antagonistic) parties are able to influence regulation, they are not able to control it or capture regulatory reform.

Another dominant proportion of research concentrates on the explanation of the reasons behind regulation in certain industries and the slower rate (or non-existence) of regulation in others. The work of Pelzman was among the first trying to explain this question, and Tollinson offered a conspiracy explanation behind vote seeking politicians’ actions to regulate natural monopolies and competitive industries and little or no regulatory activity on other industries like that of automobiles. On these lines, Becker (1978) pointed out that regulatory activity can be explained by the forces behind the demand and supply of regulatory outcomes.
The notions of productivity and regulation should not be seen as distinctive or unrelated. As regulators strive to correct market imperfections or boost the economic activity so that it harvests all the potential benefits from operations, achieving these targets could have a significant impact on the productivity and prosperity of a nation. Ceteris paribus, the structure, aims and targets as well as outcomes from regulation have a strong impact on productivity. As laid down by Hahn (2000), the selection of the right goals for regulation should be based on, among other things, the benefits and costs of each policy. Economic theory recognizes productive efficiency (i.e. the maximization of output relative to input, or equivalently, the minimization of production costs), and allocative efficiency (i.e. that subject to externalities and information deficits, economic agents-organizations-in competitive markets will be driven by their motive to maximize profits in order to generate allocative efficiency) as the two main economic goals that market agents (firms) and state principals (regulators) pursue.

Several models of economic and social regulation have emerged to address and achieve these goals. Under the public ownership model, the state protects and maximizes the public interest through direct proprietary powers over resources. This justification lies mainly on the grounds of the existence of natural monopolies (i.e. which exists when it is cheaper for production to be undertaken by one firm, rather than by several or many) and externalities (i.e. the social costs that, if negative, are not reflected in the price of the product or service, such as the escape of radioactive material from a nuclear-powered electricity generation station or the supply of public goods that individuals may benefit without paying since they cannot viably be excluded from consumption) and to a lesser degree on the notions of planning, ideology, or power. According to Wolfe (1999), examples in the UK can be found on the transfer of utilities such as gas and electricity from one form of public ownership to another due to the significant economies of scale, and the regulation of public transport in order to coordinate all different means available to the public and to remedy the severe under-investment and inefficiency, like in the case of railways.

In addition to the amount of research attempting to explain the emergence and use of regulation under the different models, as explained above, a significant amount of research examines the actual effects of certain regulatory reforms both in the long- and short-term. This primarily evaluates the revolution of regulation in terms of its impact on national economies and the targeted industries. Such impact can be realized both in the long- and short-term and can come into different forms and shapes. As a result, besides the direct or targeted effects of any regulatory reform, there could be other, less easily identified effects, namely those indirectly triggering some change in certain parameters. Indirect effects can be either expected (or planned) by the regulators or may come as a surprise to them; in other words be unforeseen and/or unintended. We illustrate the nature and relationships among the different types of effects in Figures 1 and 2 below.

Figure 1 presents the different levels of occurrence of such effects. At the first level, we can see the direct change, which is aimed for by the regulators (state). What follows is a set of indirect effects to these or other economic factors and agents, which can either be part of the planning (foreseen) and due to be dealt with in the future (intended), indirect effects that regulators did recognize or expected their occurrence but did not intended to create at that form or magnitude (indirect, foreseen and unintended), or be both unforeseen and unintended in the sense that during planning no-one could recognize/estimate their occurrence and impact.
As a result, as shown in Figure 2 below, the overall effects of regulation can broadly be expressed by the collection of:

- **Original Plan**
- **Room for ‘error’ in planning**
- **Unavoidable**

**Figure 2. Combinations and Total effects of regulation**

Consequently, we believe that at the regulatory stage the re might be some short of risk colonization, the magnitude of which partially depends on the promptness of the state to react to either planned or unforeseen impacts, on the quality of the planning process and its ability to eliminate the extend of such unforeseen (even if unintentionally occurring) effects, and the ability of regulators to ‘perceive’ that under any behavioural context, such as that of regulation, with many parameters evolving across time and geographical space and players interacting (or even lobbying for or against it), there is always going to be some unforeseen and unintended proportion of effect(s). This may represent what many analysts broadly call error. As a result, this study argues that such an error exists and triggers some effects due to either human negligence or incapacity to rule it out (perhaps due to incomplete information or just the inability of people to rule out uncertainty when they cannot (by definition) accurately predict the future). In this context, these are the unintended and, or unforeseen effects of regulation.
Research Strategy

The literature and research on regulation, in general, is enormous and spans back to the beginning of the 19th century. Our focus is not to generally address the vast amount of research on regulation, but, instead, to summarize the proportion of this research that examines regulation’s indirect, unforeseen and, or unintended consequences. We concentrate our searches for published work in three main research databases, namely JSTOR, Web of Science and EBSCO online search database. Our selection was made on the basis of breadth of each of these databases across many disciplines and their historical archive dating back to the start of the 19th century.

Our research is focused on work published within the disciplines, and consequently selections of journals, of Economics, Business, Economic History, Finance, History, History of Science & Technology, Public Policy & Administration, and Sociology. This filters out any work on other disciplines, less relevant to our perspective, such as biomedical sciences or physics, which specialize at areas beyond the scope of this study. Notably, while JSTOR offers the users the option to perform an advanced search within user-selected disciplines and their respective journals only, searches in the other two databases were made in their broadest content, since there is no such filtering option available. Consequently, the number of hits was substantially larger for the latter (unfiltered) searches. Several keywords, presented in Table 1 below, where employed in our searches in each respective database.

<table>
<thead>
<tr>
<th>Database</th>
<th>Keywords</th>
<th>Hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>JSTOR</td>
<td>Regulation AND Unintended</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Regulation AND Unexpected</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Regulation AND Indirect</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Regulation AND Unforeseen</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Unintended</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Unexpected</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Indirect</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Unforeseen</td>
<td>0</td>
</tr>
<tr>
<td>EBSCO Host Research database</td>
<td>Regulation AND Unintended</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Regulation AND Unexpected</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td>Regulation AND Indirect</td>
<td>185</td>
</tr>
<tr>
<td></td>
<td>Regulation AND Unforeseen</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Unintended</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Unexpected</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Indirect</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Unforeseen</td>
<td>0</td>
</tr>
<tr>
<td>Web of Science</td>
<td>Regulation AND Unintended</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Regulation AND Unexpected</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>Regulation AND Indirect</td>
<td>173</td>
</tr>
<tr>
<td></td>
<td>Regulation AND Unforeseen</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Unintended</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Unexpected</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Indirect</td>
<td>295</td>
</tr>
<tr>
<td></td>
<td>Productivity AND Unforeseen</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 1 presents information with respect to the keyword search per database and the total number of hits (i.e. returned results) from each initial search. In JSTOR, we can select (filter) researches to the areas of Economics, Business, Economic History, Finance, History, History of Science & Technology, Public Policy & Administration, and Sociology. In EBSCO host we perform searches by text, without the ability to apply any filtering on the types/areas of journals. Similarly, we research the Web of Science database with direct searches from the Endnote connect path to the Social Sciences Citation Index (ISI) without filtering on the disciplines.

Table 1 indicates the limited published work in journals on the unintended and unforeseen effects across time and countries. Notably, we get a significantly large amount of hits for our EBSCO and Web of Science searches on Regulation and Productivity (AND indirect effects), respectively, which may be attributed to the lack of filtering options available in these searches, as discussed above. Following our searches, we apply a two-stage filtering process to the total hits listed in Table 1 above. At this point, it is important to note the importance of applying our judgment on selecting the literature that is relevant and within the scope of our study throughout the filtering process that follows. Firstly we examine the title and abstract of the papers and evaluate whether the papers refer to regulation and its indirect, unintended and, or unforeseen effects. Then, we review the selection of relevant papers and further read their introduction and concluding parts and follow the same selection criteria in order to filter out the possibility of being misled by the titles and abstracts, and in order to identify whether each paper offers a general discussion or one drawing from a case study. We finally complete reading all relevant papers and make a list of all those offering case studies. Consequently, we end up with some 46 case studies, which we analyze and discuss in section 4 that follows.

Case Studies on the effects of regulation

In what follows until we reach our concluding remarks, we devote out attention to the presentation and analysis of the 46 cases studies illustrating the indirect, unintended and/or unforeseen effects of regulation. For stylistic reasons and due to the large size of the full table with a summary of all cases, we present it at the end of the paper. Here, we first offer a descriptive analysis with respect to the several forms and application of regulation examined by these studies, their reference to the types of effects regulation has proven to produce, as well as their evolvement across time and geographical area or country/economy/society they were addressed towards. This is presented in section 4.1 that follows. Then we discuss the aims and effects of UK regulation across time and identify the possible links or relevance of non-UK studies to the UK context.

Taxonomy of cases: Analysis by country, industry and type of regulation

This section is categorizing the case studies offered by existing literature on the effects of regulation in several groups given the country of study, time frame covered, and the type of regulation, respectively. The motivation and incentives to begin our discussion with the analysis of the cases by country (or geographical space) lie on the importance of our study to identify the volume of research on UK regulation and by that way illustrate possible gaps that may need attention and that cry for further research. Further, those examining the effects in the long-run may be more appropriate in fully addressing possible indirect effects of regulation, especially those that can be categorized (based on our analysis in section 2) as unforeseen. Finally, we were prompted to also categorize and analyze the case studies in terms of the type of regulation they refer to in order to shed some light on possible perks or omissions across literature, with our focus being placed on the UK context.
Accordingly, Figure 3 presents the classification results for the case studies by industry, country and type of regulation, respectively. Notably Figure 3 illustrates significant differences in the research investigating the effects of regulation in different countries, and further the degree of investigation among main industries and across the types of regulation adopted. More specifically, Figure 3 (a), the category miscellaneous includes 1 case study on each of the sectors of: Oil, Fisheries, Media, Brewing, and Environment. It also includes 1 study that offers a general discussion of several reforms. and shows that the majority of case studies refer to utilities, Figure 3(b) that the mostly researched regulatory reform (in terms of its effects) is the US and Figure 3(c) that, the two main areas of regulatory change addressed in these case studies are environmental regulation and that referring to privatizations (and change of control) and price controls. Also, we notice that half of the pie in (a) is consumed by research on the areas of utilities and tobacco & alcohol, and that, in (b) the same holds for research analyzing environmental regulation, privatizations, and regulation imposing price controls to certain markets or operations.

(a) Case Studies by Industry

(b) Case Studies by country

(c) Case studies by type of regulation

Figure 3. Number of cases by Industry, country and Type of regulation
To provide a more detailed analysis on the UK cases, we analyze them by industry and type of regulation and present the results in Figure 4 (a) and 4 (b), respectively. Notably, in Figure 4(a) we list each issue raised in each of the UK case studies. Although studies on telecoms, electricity and water could be included in the Utilities (general) section, we show them separately in order to highlight which (and the extend that) specific sectors of the UK economy are researched. Consequently, the Utilities (general) section includes case studies that examine the effect of some (or multiple) regulations on more than one specific sector of the utilities. On the other hand the slice named Utilities Act includes only research done on this specific piece of regulation. The importance of representing this separately from the remaining research on utilities lies on the major impact of the Act, and our attempt not to conflate the large amount of research it has triggered, which may, somehow imply its importance.

Figure 4. UK case studies classified by industry and type of regulation

Figure 4 indicates that analysis on the indirect and unintended effects of regulation in the UK is quite spread among the different industries examined, but is limited to only four types of regulation. Notably, since the majority of research is carried out for US regulation and its effects, we may attribute the larger variety of types reported in Figure 3 as only relevant to research across the Atlantic (and a small slice to the ‘rest of the world’. The authors do not imply that the world can be divided in three parts and that only the UK and US have individual characteristics. Instead, the small number of cases that do not cover UK or US regulatory reforms drives us to use such a broad definition as ‘rest of the world’. This number comprises of 1 study on EU reforms and 1 for Belgium, 1 for China and 1 for Africa) which is somewhat surprising given the constant ‘dissatisfaction’ by the UK market participants regarding its over-regulation and its impact on their operations. Also, this may be a noisy signal of the US market investigators and researchers being more uptight about the existence of such effects as opposed to their UK counterparts.
Finally, in Figure 5 and Table 2 we present the descriptive analysis of the case studies’ time spans with respect to the oldest year examined and the total period investigated (in years), respectively. Notably, UK and US research covers the same median number of years. Of course, the small amount of UK studies compared to their US counterparts and the fact that 1 out of 8 studies is covering a long time period (29 years), limit the strength of our iterations here.

Further, in Table 2 we notice that although the databases (and consequently the journals) searched span back to the early 19th century, our oldest case studies examine the period starting in 1845, 1971 and 1978 for the US, UK and ‘Rest of the world’, respectively. If we hypothesize that the longer the time span covered by a case study, the more probable for it to uncover any possible indirect, unintended and unforeseen effects from regulation, then as shown in Table 2, the median time span is the same for those studies examining UK and US effects. However, the average number of years covered by cases (in each respective country) don’t seem to differ significantly—at least not as strikingly as the spread of the periods covered as shown in Table 2.

| Table 2 Length of time period Investigated by case studies |
|-----------------|-------------|---------|-------|-----|-------|
|                 | Mean  | Median | Max  | Min | Observations |
| UK              | 12.3  | 12     | 29   | 1   | 8     |
| US              | 15.8  | 12     | 73   | 1   | 32    |
| Rest of the World | 6    | 1      | 17   | 1   | 3     |
| Total           | 14.5  | 11     | 73   | 1   | 43    |

Figure 5. Distribution of Cases Period of Examination by Country

Overall, our descriptive analysis indicates that UK research on the indirect, unintended and, or unforeseen effects of regulation seems to be dwarfed compared to that in the US. Nevertheless, research in such an area (and especially with such a perspective) may have little to do with quantity, which should only be used as a guide to unveil the possible gaps and direct towards the possible aspects in need for investigation. However, the quality, in terms of depth of investigation and ability to identify the respective types of effects may be of more importance here. To shed some light in this respect, in the section that follows we proceed with our analytical discussion of case studies.
Analysis of effects by nature of regulation

In this section we analyze and discuss the relevance of the indirect, unintended and/or unforeseen effects of regulation on productivity drawing from the context of the collection of case studies available within the large amount of existing literature we reviewed. Our basis for discussion refers to the main themes of regulatory change in relation to productivity, namely controls related to the prices (of goods or services), to quantity, the entry and exit of firms or agents in an economy and other, more general impositions, such as those related to the quality, advertising and investment. For a detailed analysis on the economic approaches to regulation, see, among others, Viscusi, Vernon, and Harrington (2005); chapters 10 and 16.

We notice that many of our case studies analyze the effects of regulation targeting or imposing changes on quality (13/43 or 30%) and quantity (10/43 or 23%). We find that 6 out of the 43 case studies (or 14%) that we reviewed and were able to assign to one of the ‘themes’ discussed above are on price and entry/exit controls by regulation, respectively. The remaining of these case studies refer to restrictions or impositions of new investments for firms (2/43 or 5%), and other, more general reforms (such as social change and protection or advertising), which account for another 14% or our sample case studies. Notably, in some cases the authors examine overlapping themes. More specifically, two cases examine the effects from regulation on quality and investment, and in one each of the pairs of entry/exit and price, quantity and quality, and quantity and entry/exit, respectively. These studies represent around 12% of case studies in our current sample.

Within these case studies, we came across the use of the term unintended effect(s) in some 11 cases (or 26%) and the term indirect effect(s) in 7 cases (or 16%). Notably, only Wolfe (1999) uses the term unforeseen effect(s), while Schwert (1997), Van Donge (2002) and Ying and Shin (1993) use as an alternative the term unexpected in their studies. We believe that this term is used to substitute for unintended effects because from the context of all three studies it is clear that the authors refer to changes that were triggered by the regulation and, as they imply, were impossible to have forecasted prior to its imposition or during the planning and implementation periods of regulation. Interestingly, we find that many authors refrain from adopting any of these terms but use other terms with equivalent meanings. For example, we come across the term ‘undesirable’ in one study; ‘further consequences’ in another; and ‘effects on the opposite direction of the goal’ in a third. Mayo and Flynn (1988) talk about the ‘dual and offsetting effects of’ regulation and Thomas (1990) adopt the term ‘differential impact’ to illustrate the unintended nature of effects. This may be attributed to the lack of a precise definition for these concepts across literature, and what seems to be the lack of an established or clear association of the collection of such effects with regulation.

Also, when reviewing the case studies collectively we come across some other main characteristics from the analysis of indirect, unintended and or unforeseen effects. Around 5% of the studies imply that when the regulation is imposed to a (highly) regulated industry (or market), adapting to the new regime it imposes seems to influence how profound the generated side effects on productivity will be. More specifically, while Berman and Bui (2001) find the existence of prior (heavy) regulation in the US oil refining industry to favour productivity increases from new regulation, Caves, Christensen and Swanson (1981) find that new regulation for less (compared to the US) regulated Canadian railroads lead to higher productivity growth (both directly and indirectly). Some 8 cases (or 19%) imply that the investigated (other than direct) effects from regulation resulted from intense reactions or even lobbying from market participants or interest groups (favourably or in contrast to the proposed changes). Only in one case (regarding the privatization of BT in the UK) we were
able to identify that indirect, unintended and unforeseen effects resulted to the alteration of the specific regulation and the re-examination of a new rule after a period of time, which interestingly could imply the most “severe” adverse outcome by such effects. Finally, another 7% of the cases refers to the problems (or advances) in economic/social infrastructure lying beneath (supporting) the indirect effects from a regulatory reform. Notably, these categorizations are not conclusive but offer some general notes on commonalities among case studies. We expect that by analysing further the goals and impacts of regulation on the basis of the case studies we may be able to identify other or more precise categories/commonalities and of course, also unveil the differences among the cases that fall within each of these categories. For more details, see WWF (2002), ‘Oil Palm Plantations

Somewhat surprisingly, apart from the work of Hahn et al. (1991) that we discuss in the introduction of the overall study, we are not able to identify any other cases referring to macroeconomic effects of regulation within our (current unfinished) sample of case studies. However, in two studies from OECD (2000) and Hills and Michalis (2000), respectively, the authors examine cross-country effects (mainly in the EU context) and Caves et al. (1981) analyze the effects by a certain regulation on railroads and offer a comparison between the productivity outcome following these reforms in the US and Canada. The rest of the studies mainly concentrate on the discussion and analysis of the (direct and indirect) impact of regulation on a specific economy/society/country.

We also notice that there are certain under investigated issues emerging from the existence of indirect, unintended and or unforeseen effects of regulation on productivity. For example, while Williams (2004) investigates the effects of an EU Act in 2000 (EU Ozone Depleting Substances Regulation in the UK) forbidding/controlling the exports of certain goods outside the European Union member states by all members, and finds that this had indirectly and unintentionally caused the piling up of used (second-hand) fridges in UK ports (that in the past were sold to third world countries) since there was no infrastructure for their appropriate recycling, the study does not examine how such restrictions affected the importing countries. This would offer more insights on the indirect, unintended and or unforeseen effects of regulation on a broader geographical space, and perhaps of a profoundly macroeconomic nature. Likewise, recent press reports show how the sharp increase in demand for palm oil in Europe as a substitute for other substances following some environmental concerns and in attempt to reduce CO2 emissions, has led to massive deforestation in Indonesia, and as a result, somehow undoing part (if not all) of the CO2 emissions prevented in Europe by releasing biomass into the atmosphere in Borneo. For more details, see WWF (2002), ‘Oil Palm Plantations and Deforestation in Indonesia. What role do Europe and Germany play’, found at [http://www.wwf.or.jp/activity/forest/lib/oilpalmindonesia.pdf](http://www.wwf.or.jp/activity/forest/lib/oilpalmindonesia.pdf), or similar articles on BBC News website.

The case studies show that there is a mixture of (indirect, unintended and or unforeseen) effects and outcomes from regulation and that in some cases these effects develop in a dynamic way in the sense that while one reform of effect (especially if indirect/unintended/unforeseen) may positively affect productivity another effect (or set of such effects) may significantly reduce it. To illustrate this we attempt to assess the overall impact of indirect, unintended and or unforeseen effects on productivity for each case study (and subsequently per country it refers to) and present the results in Table 3 below. Of course, since it is not quite straightforward and is difficult in principle to add positive and negative signs representing the impact on productivity- from qualitatively different cases up, this table should be interpreted with caution.
Table 3: Overall impact of indirect, unintended, unforeseen effects of regulation on Productivity, per country

<table>
<thead>
<tr>
<th>Theme</th>
<th>UK</th>
<th>Overall</th>
<th>US</th>
<th>Overall</th>
<th>Rest of the World</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Per case</td>
<td>Overall</td>
<td>Per case</td>
<td>Overall</td>
<td>Per case</td>
</tr>
<tr>
<td>Quantity</td>
<td>+</td>
<td>-</td>
<td>+/−</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>-</td>
<td>+/−</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Entry/Exit</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Quality</td>
<td>-</td>
<td>-</td>
<td>+/−</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>+/−</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>No cases</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Price controls</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>+/-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Investments</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>No cases</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Other</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>No cases</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>+/-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3 presents the impact of the effect(s) of regulation on productivity per case and overall for each theme on regulatory change and its relation to productivity measures. The Overall impact is calculated on the basis of which sign (+/-) prevails on the All effects section of the table.

Notably, only a small fraction of research concludes or implies that such side-effects are beneficiary for productivity. Also, it seems that the theme of regulation may be important in
assessing such effects. In the UK, while quantity regulation seems to have had a positive effect on productivity, other forms such as entry/exit had the opposite effect, while the case studies examining the effects of price control regulations seem to be equally split between the two opposing directions. In the US, the outcomes seem to follow similar patterns. Specifically, price controls, quality, quantity and other (such as social) regulation seems to have had a detrimental effect on productivity, whereas, entry/exit controls positively affect it, and the results are mixed for regulation aiming at shaping investment activity. Finally, the small amount of cases with data from the ‘rest of the world’ also seem to point towards a negative (or mixed and inconclusive) impact of the indirect/unintentional/unforeseen regulatory effects on productivity.

**Concluding Remarks**

This paper offered an analysis of the nature and impact of the indirect, unintended and or unforeseen effects of regulation on productivity by reviewing a collection of case studies stemming from appraising a large amount of published literature. We provide some analysis indicating that research on such consequences by regulation is not as substantial as their impact on productivity (as shown by current case studies around the globe) would imply. Moreover, our analysis indicates that this need for further research may be more intense in the context of the UK economy since only a small amount of work has been carried out, which nevertheless hint their role in the productivity slow-down associated with specific regulatory reforms over the past 30 years. Also, our results indicate that certain regulatory changes (for instance privatizations that dwarf all other governmental activities) and industries (such as utilities) are more researched than others. The latter may be explained by the hype around the privatizations over the 1980s (especially in the UK) and the concerns by researchers in evaluating the impacts and performance of such governmental policies.

The importance of this study is highlighted in the fact that it addresses and analyses the issue of indirect, unintended and or unforeseen effects of regulation on productivity in a collective manner, and by investigating existing literature (and case studies) it attempts to shed some light to the overall (or subject to interactions between them) effects. Finally, the study unveils the importance of research on this field, and particularly, that recognizing and analyzing the various forms of effects and their dynamic (if at all) relationships across time and space. This need may be more intense for the UK context, where there is a clear imbalance between the volume of regulation and the respective research and assessment of its effects both in the long- and short-term, with the first dwarfing the latter.

**References**


Burton, J. 1997. The competitive order or ordered competition?: The 'UK model' or utility regulation in theory and practice. Public Administration, 75(2): 157-188.


Performance Measurement, Regulation & UK Productivity

Kim Tan and Rosalind Rae

Abstract
The paper presents a systematic literature review on the links between regulation and performance measurement (PM) specifically focusing on the unintended consequences. The findings illustrate a number of consequences both positive and negative of regulation and performance measurement upon and within the firm. The distinction of whether these consequences are intended or unintended is made clear in only a few papers. The evidence provides a brief overview of the benefits, consequences and some unintended consequences of regulation and performance measurement. Some benefits of regulation on performance measurement are: to provide a benchmark for cross-firm comparison in an industry sector; improvement in standards; ensures control and accountability; improvement in performance of the firm. Some consequences of regulation on performance measurement are: an increase in costs to the firm through increase in measurement, recording and reporting; reduction in efficiency due to an increase in inputs needed to achieve the same output. The unintended consequences are the commodification of services and deprofessionalization of the worker; the move of the firm to focus on unregulated activities to generate shareholder wealth; a tendency to measure what’s easy not what’s useful; can get conflicting measures whereby achieving one target is to the detriment of achieving another. The literature review identifies significant gaps in the research with respect to exploring the relationship of regulation and performance measurement and identifying and providing evidence on the unintended consequences of regulation and performance measurement upon and within the firm. The study highlights the need for further research to enhance our understanding.

Introduction
The literature review aims to explore the relationship between regulation and performance measurement on a micro level (firm level) and to classify the unintended consequences of regulation on performance measurement within firms located in the UK. The objectives of the review were:
- To explore and evaluate previous work on the relationship of regulation and PM;
- To enhance our understanding of the links between performance measurement and regulation and its effect on productivity;
- To establish the nature of the relationship between performance measurement and regulation;
- To explore examples in the literature on the unintended consequences of performance measurement and regulation;
- To identify areas for future research;
- To recommend areas for future research.

For the purpose of this research the focus was centred on what the ‘unintended’ consequences are. Using the dictionary definition of ‘unintended’ consequences, it is “an activity which has produced a consequence that was not planned with any intent or purpose”. The result is something which could have a positive or negative effect on the firm’s environment.
Within the scope of this study, the evidence found in previous research which focuses on the unintended consequences of performance measurement and regulation appears to be sparse. The research conducted has been focused on single case studies in particular industry sectors, which makes it difficult to gain a clear overview of the area of investigation.

Another challenge which previous work fails to address is the classification of the consequences of regulation and performance measurement into intended and unintended. On many occasions the differences were not made clear. The evidence provided very little information on the unintended consequences. However, research conducted by Adcroft and Willis (2005) looked at the (un)intended consequences of performance measurement in the public sector and found that the result was a commodification of services and depersonalization of the workers. Humphreys and Francis (2002) looked at airport performance measurement and found that there were significant dysfunctional effects related to performance measurement resulting from regulation and commercialization. For instance, the achievement of one performance measure had a negative impact on the achievement of another, so the two measures were at conflict with one another. They also found that as a result of heavily regulated activities, the firm focused on unregulated activities to generate wealth and to meet shareholders requirements.

Overall the research identified and explored was relatively thin on the ground with respect to focusing on the unintended consequences of regulation and performance measurement. This goes some way to the identification of a research gap and the opportunity for further exploration. The consequences of regulation and performance measurement can impact on the firm both positively and negatively.

In this paper, the findings linking regulation with performance measurement will be presented outlining the unforeseen consequences resulting in this linkage. A general overview will be
explored to enhance our understanding about what has been done previously, what the challenges are related to this area of research, and what the opportunities are.

**Methodology**

Regulation, productivity and performance measurement have been researched widely with vast amounts of information located in their respective fields. The study’s aim was to explore the literature for research on the relationship between regulation and performance measurement on a micro level and to identify the research gaps in this area with the ultimate aim of scoping a study for further research. It became apparent fairly early on in the literature review that the biggest challenge would be efficiently sifting through all the research material to locate the relevant work. The scope of the research involved a systematic review exploring aspects of the literature and empirical evidence to identify themes relating to the relationship between regulation and performance measurement on a micro level.

A number of stages make up the review of the literature and have been designed with the scope of the study in mind. Due to the relatively small scale of this particular research project with respect to the two particular areas of regulation and performance measurement, one of the challenges was trying to make priority decisions on what was relevant and what was not relevant in a time efficient manner. This effectively meant making decisions early on to identify key research papers and drill down to the detail earlier than perhaps one would necessarily do so. This resulted in making decisions on the content of the abstracts as to whether the journal merited further exploration. It soon became apparent that a small amount of research had been carried out specifically looking at the unintended consequences of regulation and performance measurement on a micro level. The search also involved sifting through related materials which might have had links but it wasn’t made clear in the abstract or it appeared not to be the research’s primary aim.

The research steps were as follows:

- Keywords were identified through brainstorming the areas of interest. For example, performance measurement, performance assessment, the consequences of regulation, regulation and performance evaluation are just some examples. Details are outlined in Appendix 1 of keyword searches;
- The keywords were used on their own and as part of a more advanced search connected with other key words in database searches. For example, regulation* AND performance measurement*. The literature search was carried out as a subject search criteria initially and if zero items were found then as general text search within the databases. (This was only done when there were very few items located when the subject context was used, and then to assess the likely relevance of other items found under the wider domain);
- Key words were used to search within 8 search engines. These included EBSCO Business Source Premier, ABI Proquest, Web of Science and specific operations management journal databases (listed in Appendix 1);
- Once the articles were identified which matched the search criteria, the abstract was read and evaluated by the researcher deciding whether the article fitted into one of three categories, category ‘A’ for being of direct relevance, category ‘B’ for being vaguely relevant but not directly and category ‘C’ for being interesting but not really relevant for this particular study;
- The list of ‘A’ articles were read and evaluated as to whether they could contribute to the study in question.
**Methodological Challenges**

The limitations of the approach were as follows:

- The systematic review was not as thorough and as wide-reaching as one would have wished. The scale of the review was limited by the project scope and it would have been desirable to have been in a position to analyze the literature materials using Procite and Nvivo to enable a larger number of journals to be assessed early on in the review process. All the literature was reviewed by the researcher manually and assessed by reading the individual abstracts at the beginning of the evaluation process. The volume of the papers located was vast and it was a huge challenge to identify efficiently and effectively which would be relevant. The scope of the project determined the volume of papers scrutinized and subsequently reviewed;

- A limitation to using this method would be that if the abstract did not explain categorically what the article was trying to achieve and what the findings were then there was a possibility it might have been missed;

The industries studied in the literature were focused mainly on the Public Utilities, with water being a particular favourite because these areas have been the most researched with respect to being regulated industries. This makes any generalizations outside of those particular industries quite difficult. There was also a bias to focusing on Development Envelope Analysis (DEA) as a performance measurement technique and how firms use this to measure their performance. Our aim in the study was to get much more of an overview of what work had been carried out previously and to explore the links present, not to focus on one particular measurement technique and one particular firm located in a particular industry. It was hoped that this approach would enable us to form much more of a helicopter view of the general consequences including positive and negative and the unintended consequences of what had been researched previously.

All the material evaluated was from companies based in the United Kingdom. This was chosen as a good starting point and base for further research. From the initial findings of the literature search, a conscious decision was made to focus the research on the links and consequences of regulation and performance measurement because there was not enough material located to be able to concentrate on just the unintended consequences (only a handful of journals had this as their primary focus), so it was decided to open the search up further to ascertain what the themes were coming out of the related journals and to see whether any relevant material could be located in these articles.

The majority of the literature located has been very specific in its focus and has been carried out on a case-by-case basis. This in itself provides challenges and opportunities. This approach enables specific firms within one industry to be studied in detail which can provide valuable insights into the thoughts, behaviour, decisions and processes that are present in that particular firm and related to that particular industry. However, what this approach doesn’t do is enable any kind of generalizations to be made.

An interesting find of the search was the lack of presence of relevant material in key operations management journals in covering links of performance measurement and regulation. Performance measurement is a huge field in Operations Management literature but when trying to explore links with regulation; the material is somewhat thin on the ground.
The literature review sought to investigate the existence of the following propositions:

1. Regulation impacts on the performance measurement of the firm, from the strategic level to the individual employee level. Regulation can have positive and negative implications for the firm. Some of the outcomes of regulation and performance measurement are intended and expected but relatively little is known regarding the unintended consequences of regulation and performance measurement at firm-level.

2. The relationship of regulation and performance measurement is an area relatively unexplored. Strong links exist between regulation and performance measurement at firm level, in which regulation can have significant impact on the performance of the firm and the performance measures it uses. It might also be argued that the outcomes of performance measures can influence regulation.

3. Regulation and performance measurement can affect the firm in different ways. The impact can be dependent on firm size, the environmental context the firm sits in and the social context the firm is in. In addition to this, the short-term implications of regulation and performance measurement on the firm can be very different when comparing with the long-term implications. For instance, in the short-term, regulation and performance measurement increases costs and reduces efficiency for large and small firms. The impact of increased costs and reduced efficiency is much greater for the smaller firm. Larger firms may see no significant difference. However, it is argued that in the longer-term, firms will attempt to find ways to reduce the costs and increase efficiency (reduce the impact of regulation) through process improvement and innovation but this will be achieved quicker by the bigger firm who has more resources to dedicate to this.
The literature review identified several articles which actually mentioned unintended consequences of performance measurement and regulation. Only a few of these journals focused on the unintended consequences as the main focus for research, the others just mentioned the presence of unintended consequences but it was not the focus of the research of the journal. However, a number of articles have made some interesting points relating to the consequences of regulation and performance measurement. On the whole, the consequences were not clearly defined in so far as to say whether they were intended or unintended. Some consequences of regulation are obviously intended some appear to be unintended. Table .1. Summarizes the key authors and articles.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Date</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adcroft &amp; Willis</td>
<td>2005</td>
<td>Looked into the (un)intended outcome of public sector performance measurement, in particular looking at examples from the NHS and the education sector. They found that an increased usage of performance measurement techniques in the public sector resulted in the commodification of services which was delivered by an increasingly depprofessionalised public sector workforce</td>
</tr>
<tr>
<td>Brigham Fitzgerald</td>
<td>2001</td>
<td>Analyze the relationship between individual &amp; organizational performance management &amp; measurement within a regulated water company. They propose 4 dimensions of control constitute the social relations of economic regulation: mediation &amp; negotiation; visibility of reporting; prioritization of performance measures and perception of control</td>
</tr>
<tr>
<td>Humphreys &amp; Francis</td>
<td>2002</td>
<td>Looked at the past, present and future of airport performance measurement &amp; focused on the changing ownership of airports from public to private interests on performance measurement systems. They found that measurement systems were developed in response to changing organizational contexts. They concluded that airport performance measures are important for day to day business and operational management, regulatory bodies, government and other stakeholders such as passengers and airlines</td>
</tr>
<tr>
<td>Shaffer</td>
<td>1995</td>
<td>Focuses on the consequences of public policies for the competitive environment of the firm. He stated that firm level responses can be strategic adaptation and attempts to influence policy. Organizations protect and advance their political interests through environmental scanning, lobbying, political actions committees, coalition building (like trade associations) and advocacy advertising</td>
</tr>
</tbody>
</table>

The key points of the literature review are split into the following sections:
- Introduction to the relevant literature;
- Some Benefits of PM;
- Some Weaknesses of PM;
- Change in Management Behaviour as a consequence of Regulation and PM;
- Unintended outcomes;
- Research Methods.

**Introduction to the relevant literature**
Adcroft and Willis (2005) looked into the (un)intended outcome of public sector performance measurement, in particular looking at examples from the NHS and the education sector. They found that an increased usage of performance measurement techniques in the public sector resulted in the commodification of services which was delivered by an increasingly deprofessionalised public sector workforce. It was found that there was a real push to import private sector performance management techniques into the public sector and there was much skepticism about how much private sector management practices had transformed the way in which the NHS operates and is managed. One of the reasons they saw this lack of improvement was down to the differences between public and private sectors. For example, some of the technical and managerial issues that exist with standard public sector performance measurement systems meant that the performance measurement systems were not fit for purpose. There were difficulties in importing management practices from one context to another, in this case from private to the public sector. The measurement systems had resulted in a strong focus on outputs and results, which had resulted in more privatization, contracting out and internal market reforms in the NHS. There were preoccupations with performance measurements which were built on the assumption that it would bring real and tangible benefits to organizations. However, in a complex organization, performance is determined by a whole series of internal and external factors which complicates the measurement of the right output.

Humphreys and Francis (2002) looked at the past, present and future of airport performance measurement and also at the changing nature of the performance measurement of airports. Performance measures were primarily used to assist in making publicly owned airports
accountable to their government owners. The focus was placed on measuring the performance of airports and their roles within the wider air transport system. Examples of different purposes to which performance data may be used:

- Government – for economic and environmental regulation;
- Airline – for cost/performance comparison across airports;
- Airport Managers – to run their business;
- Passengers – to assess how well they are served as consumers;
- Owners/shareholders – to assess business performance and the return on their investment.

They focused on the changing ownership of airports from public to private interests on performance measurement systems. They found that measurement systems were developed in response to changing organizational contexts. With pressure for change coming from changing ownership patterns; an increased commercial focus; regulation; rapid passenger growth; increased concern for the natural environment and technical innovation. The challenge for establishing an appropriate performance measurement system (PMS) for an airport is huge – there are many interacting parts: passengers; airlines; handling agents; surface transport service providers and the interests if regional and the national economy. They found that PM is a critical management activity at the operational level of the individual airport and at the wider system level. Airport performance is measured for efficiency from a financial and operational perspective (Doganis, 1992); to evaluate alternative investment strategies; to monitor airport activity from a safety perspective and to monitor environmental impact.

Humphreys and Francis (2002) concluded that airport performance measures are important for day to day business and operational management, regulatory bodies, government and other stakeholders such as passengers and airlines.

**Some of Benefits of Performance Measurement (PM)**

Meyer (1994) suggested that performance measurement is useful in so far as it can ‘tell an organization where it stands in its efforts to achieve goals’. Drucker (1995) suggested that PM generated new and additional resources; clearer understanding of economic chains and generated wealth. Kaplan & Norton (1992) in discussing the balanced scorecard, they stated that the benefits were in the translation of the ‘company’s strategy and mission statement into specific goals and measures’, it allows for products to get to market sooner and innovative products are tailored to customer needs. Ness and Cucuzza (1995) went on to say that the average quality of decisions made day in day out will be vastly higher than before. Other benefits outlined were:

- To reduce variations in performance and service provided;
- To share best practice;
- Will focus management of key aims like quality and effectiveness;
- Ensures control and accountability;
- A way of evaluating performance;
- Motivates individuals to meet certain performance targets;
- Improvement in standards;
- PM provides a consistent and coherent process for delivering the goals and strategy of the organization. The process is centered on tangible and measurable outcomes and provides;
Adcroft and Willis (2005) found that from a national perspective performance measurement has allowed comparison across hospitals and health care trusts in terms of their performance whether it is good or bad, the aim being to raise standards, focusing on certain activities and outputs through increased transparency, more accountability and to create competition between hospitals and trusts. As a reward, hospitals that performed well attracted more funding.

Humphreys and Francis (2002) found that management required information to enable them to identify areas that are performing well and those where appropriate corrective action needs to take place. Governments typically require information to regulate airport activities. The airports customers will also be interested in assessing its performance. Airports have recognized the potential for benchmarking against other airports to improve their competitive position through the identification and adoption of best practices. IATA produces a comprehensive comparison of airport service performance indicators from the perspectives of international passengers. These ratings offer a starting point from which managers can start to ask questions about performance levels. Humphreys and Francis postulate that the greater the extent to which the airport starts to analyze the processes that generate these figures and learn from best practice is likely to advance as a consequence of the drive towards a more commercial business focus.

Shaffer (1995) outlined some competitive gains for businesses resulting from public policies outlined as:

- Increasing total market size;
- Advantages over rivals;
- Reducing the threat of rivals and product substitutes;
- Gaining bargaining power over suppliers and customers.

(Gale & Bulchholz, 1987; Cory & McWilliams, 1994).

Some Weaknesses of PM

Meyer (1994) points out that PM is less useful in explaining what the organization should do differently. Adcroft and Willis (2005) argued that PM doesn’t take into account the softer measures and therefore it is more difficult to quantify which measures still have influence on the performance and delivery of the goals. Performance can be affected by external factors as well as internal factors like market conditions, industry structures and social settlements. They found that individual elements measured are rarely independent of one another; therefore they pose the question, does it really give a true measure?

Other weaknesses have been listed as:

- Improving performance in one area may result in worsening performance in another (especially if resource is moved);
- Scientific approaches to measurement assume objectivity achieved through dispassionate analysis of the available evidence, but subjective interpretation of the evidence is often the case;
- It is impossible to get a 100% objective measurement;
- The use of results on a national scale for comparison purposes (like league tables) assumes everybody starts from the same point. There are usually more sources of difference between same-service providers than there are similarities (some will have more resources than others, some hospitals will be located in more affluent areas, league tables do not reflect these differences);
- In any complex process of service delivery there will always be elements of that service which are beyond scientific measurement. In such cases, proxies are used whose relationship to the thing being measured can often be tenuous;
The choice of targets and performance measurements can be used to cynically influence the results of measurement. The choice of what and how something is measured can often be about creating an impression of improvement rather than delivering any real improvement. E.g. measurement of crime – measure the ones that are easier to solve (Adcroft and Willis, 2005);

Humphreys and Francis (2002) found that airport operational performance measures that relate passenger level of service to international standards are still widely used. The major weakness for this kind of measure is that it is too crude. As Gosling (1999) highlights there is a tension in data collection between what is easy to measure and what is useful to measure but potentially more difficult. The level of service delivered is contingent upon various passenger characteristics and a certain design may deliver totally different levels of service for passengers depending on the purpose and the nature of the journey.

**Change in Management Behaviour as a consequence of PM**

Shaffer (1995) focuses on the consequences of public policies for the competitive environment of the firm. He stated that firm level responses can be strategic adaptation and attempts to influence policy. Organizations protect and advance their political interests through environmental scanning, lobbying, political actions committees, coalition building (like trade associations) and advocacy advertising. Many managers view attention to governmental affairs as a vital part of their jobs both as a defence against regulatory intrusions and as a means of gaining corporate advantage. Firms attempt to control the political agenda for competitive gain.

Shaffer (1995) stated that the sources of literature are diverse and said ‘scholars interested in business political activity is to a large extent fragmented by their disciplinary loyalties’. As a consequence of this, research in the general area of business political behaviour has proceeded unevenly, and is characterized by a lack of theoretical and methodological cohesion. Firms may support legislation and regulation that benefit their position vis-à-vis rivals, entrants, substitute products, buyers and suppliers. This is termed the strategic use of public policy (Mitnick 1981; Wood 1985), for the purpose of gaining competitive advantage.

Oster (1982) put forward the concept of the formation of industry sub-groups that firms use regulatory processes to hurt rivals, even when the collective welfare of the industry is damaged:

“Indeed, a firm may even encourage passage of regulation which reduces industry demand or increase costs. The firm may encourage such regulations because they differentially damage its rivals, and thus rearrange market shares at the same time they reduce the total market”.

Adcroft and Willis (2005) stated that managers had options to choose as a response to the interpretation of the measures and targets that had been set:

- Strategic processes are drawn out from the balanced scorecard (Kaplan & Norton 1992);
- Improved management of teams (Meyer 1994);
- Return on management activities (Simons & Davila 1998);
- ABC activities (Ness & Cucuzza 1995);
- The performance management manifesto (Eccles 1991);
Humphreys and Francis (2002) found that airports whilst moving towards privatization and the new commercial emphasis at European airports led to new performance measures being introduced to reflect changing managing goals. New measures fall into three categories:

- Financial measures to monitor commercial performance;
- Measures to meet the requirements of government regulators;
- Environmental measures.

Commercial pressure from ownership forms that demand a degree of financial accountability have led many more airports to focus on measuring operational and business performance within the airport company.

Airports are much more financially orientated, and these measures reflect the diversification of the business under new commercial and privatized ownership structures and subsequent management drive to satisfy shareholders.

(Un) intended outcomes

The Hawthorne Experiment is a classic case study which illustrates a number of things. Gillespie (1991) surmised that ‘the most striking characteristic of the Hawthorne experiments were the comfortable assumptions which had been formed pre-experiment, which were ‘dramatically overturned by the force of unexpected and irrefutable experimental evidence’. In other words, they had produced unintended consequences/results’ through the experiment which they had not anticipated or foreseen beforehand. There were a number of experiments carried out at the Hawthorne plant. The first was to prove that improvements in the working environment; a change to the level of lighting provided, for example, would improve the productivity of the workforce. An increase in productivity was seen but could not be isolated to just being down to the lighting, as when the conditions were returned to their original state, productivity stayed at the new higher level.

The Hawthorne experiments seemed to show that the managers had to look beyond the technical organization of the factory, as their productivity was influenced as much by personal attitudes and informal social organization, as they were by the formal lines of organization and authority. The important message from the research conducted by Mayo and Barnard in the 1930’s was that it was the attention to the employees, not work conditions per se, that had the dominant impact on worker productivity. In fact, this effect is now known as the Hawthorne effect, to refer to unexpected influences of non-experimental variables in any experiment.

Adcroft and Willis (2005) found the unintended outcomes were:

- The increased use of performance measurement and the importation of private-sector management principles and practices will have the dual effect of commodifying services and deprofessionalising public sector workers;
- Commodification is seen as transformation of relationships into quasi-commercial relationships with an emphasis on economic activity of buying and selling and the management activity of performance measurement;
- Deprofessionalisation would be the result of ‘worth into exchange value’ and the conversion of the highly skilled knowledge worker into paid wage labourers;
- Transformation through commodification changes the basis of decision-making such that values become much less important than the rules, regulations and performance measures of the organization.
Meyer (1994) concluded that ‘the long held view of what gets measured gets done has spurred managers to react to intensifying competition by piling more and more measures on their operations in a bid to encourage employees to work harder. As a result, team members end up spending too much time collecting data and monitoring their activities and not enough time managing’.

An unintended outcome found by Humphreys and Francis (2002) was the difficulty in producing a PM which was useful at the right levels. Airports are complex organizations and what their study illustrated was that a PM produced at one level was often of little use through-out the rest of the organization. For example, annual airport performance results were collated in the United Kingdom on a national basis and made available to the public. The measure of output was a key performance indicator (passenger and freight). However, a problem here was that resources used to generate the output for passengers and freight was very different. Differences included: the ground handling equipment used; the hours of operation; the average age of aircraft; and associated environmental impacts. These different inputs and outputs required measurement in their own right. There was little point in amalgamating the two other than to get an overall financial result. The WLU was accepted for a measure as a public utility view but was little use to airport management.

Another unintended outcome found by Humphreys and Francis (2002) was that regulation of airport activity by the government and the performance measure’s that accompany it can have dysfunctional effects. The regulation introduced as part of the UK’s privatization policy and the impacts of this on the environment are in conflict. BAA view their retail activity as a core competency and have purposefully diversified their efforts into increased airport retail activity. The commercial pressure for increased retail activity (due to a price cap imposed by the Government on the level of aeronautical charges) within BAA’s terminals may conflict with the United Kingdom policy to maximize the use of airport capacity. The need for a new terminal may be reduced if the retail space was used for passenger processing. The BAA has also diversified into property development, and has purchased and managed a chain of hotels. In addition, they have entered a joint venture to develop and manage out of town shopping centre’s as part of their strategy to maximize their return to shareholders by focusing on unregulated activities.

Another dysfunctional effect of performance measures – is when one measure negatively impacts on another. For example, Humphreys and Francis (2002) found that aircraft noise restrictions were imposed in an operational day to protect the local community, but caused stacking, which had a negative impact on the environment through increased levels of emissions from queuing traffic. It is much more difficult to measure the consequences of emissions. Performance measurements need monitoring in order to identify and correct such dysfunctional effects. In addition, due to environmental restrictions which restrict the take-off and landings on runways, runways operate less efficiently and often at the expense of environmental pollution.

Also, the drive to raise commercial income levels may conflict with environmental goals by generating more surface access trips from non-passengers wishing to access the retail activities. There is a conflict between meeting shareholder value and environmental targets.

Shaffer (1995) found a number of authors outlining some unintended consequences. Mizruchi & Koening (1986) and Shaffer (1983) looked at the bargaining power of vertically related firms with opposing political and economic interests. They found that government
intervention may enhance the relative position of one party at the expense of the other. Pashigan (1994) found that environmental regulation favoured the survival of the large plants at the expense of the small plants. Dean and Brown (1995) showed that rules for compliance for pollution regulation lead to capital requirements that discourage new firms from entering the market. These studies suggest that firms adapt to regulatory incentives and they do not suggest that large firms intentionally sought regulations to deter small firms. Russo (1992) showed that electric utilities diversified and vertically integrated into unregulated business sectors in response to increasingly hostile regulatory constraints.

Regulation can have asymmetric affects on competing firms (Leone 1996). As a result firms with superior capability for adapting to regulatory dictates may also attain a position of competitive advantage over their rivals.

Knight et al (2005) compared the manufacture of and source of materials in reinforced wood doors and steel doors. The focus was on the environmental impact of the two processes. The aim of producing more steel doors and reducing the production of wood doors was to reduce the usage of wood and therefore achieve positive environmental results through saving forests, reducing carbon dioxide emissions, decreasing soil erosion, to name a few. However, an unintended consequence of trying to achieve one set of environmental targets, negatively impacted on others through the increase in energy needed to manufacture the steel doors.

McKinney and Miller (1998) looked at the results of a survey completed by over 200 manufacturing firms in the USA, to determine the effects of compliance with environmental regulations on manufacturing operations. They found that many companies did not achieve full compliance with the environmental regulations. Some of the challenges facing the companies were:

- Staying up-to-date with the huge volume of regulations;
- The complexity and confusing nature of the regulations themselves;
- Most significantly for this research they found that compliance with some regulations were in direct conflict with the achievement of other regulations.

They concluded that environmental regulations were difficult to implement and that more funding was consistently needed for the maintenance of and future changes to regulation.

Research Methods

Shaffer (1995) found that empirical research linking firm-level strategies to political activity appear to lag behind the production of conceptual papers. Empirical approaches range from large scale statistical analyses, utilizing a variety of secondary data sources, to case studies of one firm or a small sample of firms. Case studies being the favoured research approach due to the highly situational nature of the strategic-political issues involved and the varied attributes of different industries and regulations (Mahon, 1983).

As a research problem, political influence activities may be hard to describe using statistical methods. Events are unpredictable and sporadic, outcomes are hard to specify, coalitions are transient and the environment is extremely complex. The qualitative nature of the case study approach emphasizes the description of complex organizational processes and interdependencies. The limitation was the problem of generalisability.

Highly quantitative methodologies such DEA and TFP have been applied to airports in order to measure inputs in relation to outputs. Other studies have used a one case study approach which has been industry and firm specific.
Theoretical Perspectives

The studies have used different theoretical concepts underpinning the research. For example, Brigham and Fitzgerald (2001) used the principal-agent concepts. Shaffer (1995) specified that each of the studies he looked at used contingency and resource dependency theories arguing that organizations adapt to environmental uncertainty by a process of internal realignment. Ungson et al. (1985) stated that “the eventual effectiveness of organizations depends largely on the degree to which they are able to adjust and modify internal structures and processes to accommodate the requirements of these regulatory sectors”.

Conclusion

Some unintended consequences of regulation and performance measures have been identified through the literature review. The existence of unintended consequences has been proven (Proposition 1). A significant research gap exists in exploring in more depth the scale of the unintended consequences and what could be done to account for these. It is believed that there are strong links between regulation and performance measurement (Proposition 2). However, whilst this literature review found evidence of links between regulation and performance measurement, they were far from being exhaustive in nature and could not be quantified through this particular study.

Regulation and performance measurement can impact on firms in different ways depending on certain characteristics. Also critical in looking at this is assessing the short-term implications versus the long-term impact on the firm. It is believed that the consequences for larger firms will be significantly different to smaller firms (Proposition 3). The review found limited evidence to support this, however, it is believed to be an important area for future research in terms of developing best practice and understanding the impact on different firms, and why this difference exists.

The scope of the study was limited and as a result the data collected on the benefits, weaknesses, and unintended consequences have been far from exhaustive. The review has found literature which has had a focus on the more negative effects of regulation and performance measurement. The data explored has suggested that the area is not well researched especially when looking at the unintended consequences of regulation and performance measurement at a micro level and consequently, its effect on productivity. The research has been industry and firm specific and there is no clear classification between the intended and the unintended consequences. The scope for extending the literature review and for conducting further research in this field is large.

The review of the evidence concerning the links between performance measurement and regulation especially the unintended consequences highlights the need for some further extensive literature review and research. The evidence has provided confirmation of the presence of a substantial research gap exists when looking at the links between regulation and performance measurement and especially the identification of unintended consequences of this relationship. The focus of previous research has been firm and industry specific, with a few journals specifically focusing on the unintended consequences. The research for regulation and performance measurement are vast in their own fields but little evidence was found on the relationship between these two and their link and impact on productivity, especially from an operations management perspective. Another area for future research is the classification and clarification of intended and unintended consequences.
Further research will contribute significant knowledge to an emerging, and yet important area in productivity related research. A good understanding of the linkages between performance measurement and regulation will assist Policy Makers by providing knowledge about the unintended consequences and how policy might take these into consideration when setting policy frameworks by minimizing the negative effects and being able to take advantage of the positive consequences.

A good understanding of the linkages between performance measurement and regulation could also assist firms and the managers within them to understand the issues that arise and to try and design performance measures which are realistic, useful and satisfy a number of stakeholders. The issues raised from conflicting performance measures can then be understood and minimized.

Further research would seek additional evidence for the propositions made. Research would look at how companies respond to regulatory requirements and how this was translated into performance measures through-out the different levels within the firm. In the first instance, a case study approach would be used to explore the relationship between regulation and performance measurement and the unintended consequences in detail and the study would be performed on multiple levels to allow data to be gathered from different perspectives. A heavy regulated industry sector would be used for the study to try and flesh out some of the themes and issues that exist.

References
‘Regulatory Design: External and Internal Performance Measures in Regulated Industries’, ESRC project number RO22250147.
www.dti.gov.uk/ewt/betterregulation.htm
www.cbi.org.uk
www.businesslink.gov.uk
Review of Indirect, Unintended and Unforeseen Consequences of Environmental Legislations: A Preliminary Investigation

Dr. S. J. Antony and Dr. Anjula Gurtoo

Introduction

The environmental impact of products and practices has become an important issue of debate and concern over the past few decades in United Kingdom (UK). Through a combination of public pressure and government intervention through legislations, stakeholders (companies, institutions, consumers and others) are being forced to consider the environmental impact of their actions. The main philosophy behind environmental policies has been the polluter pays principle which was initially incorporated by the EU in the EEC treaty of 1986 and later adopted by UK. The concept of producer responsibility has grown out of this principle. However vastly different interpretations of what constitutes a ‘producer’ responsibility has led to the current status of ‘shared responsibility’ which divides the cost among various stakeholders and sectors of the industry.

The purpose of environment legislations is to prevent and reduce environmental problems caused by production, use, or disposal of goods. Some of the common features of these legislations are phasing out of certain substances; producer responsibilities in production, waste collection and recovery of the product, and duty to provide information and documentation to the government and consumer; government approval for treatment of certain substances; and end of life clause for certain products. The direct impact of this legislative control as identified in literature include product and process redesign to include environmentally friendly materials and processes; life cycle management for waste disposal, recycling and reuse; and government structures and mechanisms to monitor compliance (Low and Williams, 1998; Kim, 2002; Hug, 2001 and many others). These changes in turn have had some direct consequences on various stakeholders like increased costs and responsibilities, and change in nature and focus of R&D and innovation (Kim, 2002). However, these legislations, as in case of other legislations, also have the potential to produce secondary and tertiary impacts on various stakeholders which are either not clearly visible or take time to show their impact.

Several studies highlight the significant impact of regulations on productivity and the report of ESRC [2004] clearly points at environment legislations as one of the possible sources of the UK productivity gap. However, only few studies have attempted to understand the relationship between the secondary or tertiary impacts of these regulations, and productivity and its related areas. If one casts a net wide enough however by defining ‘impact or consequence’ rather broadly and search indirect as well as direct evidence it is possible to identify some studies potentially capable of shedding some light on this cause and effect relationship. This paper is a systematic review of these studies in order to identify secondary and tertiary impacts of environmental legislations on various stakeholders with special search for, End-of-life (EOL) legislation, with the objective of understanding possible reasons for the UK productivity gap.
Along with the challenge of identifying the nature of these tertiary or secondary effects, on various societal stakeholders including the decision makers themselves, i.e., the government, defining these effects has been another challenge. The secondary and tertiary effects are classified as indirect, unintended and unforeseen according to the nature of effect it is seen to have on one or more stakeholders.

**Defining direct and not so direct**

There is a growing understanding that environment factors have an impact on society (Rosenzweig et al., 2001; Griffin et al., 2001; Deudney, 1999). There are direct cause and effect relations between environment and society that do not require debate. However identifying indirect causality, complex causal chains that bring unpredictable surprises and the reflex nature of the environment requires creative analysis and is a methodological as well as scholarly challenge (Hug, 2001). A simple human management decision may lead to changes in the environment which in turn can impact human population in new and often unforeseen ways. For example, a simple governmental decision of forest fire in Indonesia to clear land for agriculture caused a cloud of smoke to cover much of South Asia (Fraser et al., 2003). Similarly, regulatory facilitation of intense industrial activity of a certain kind in Canada and Australia, which created clouds of aerosols, is linked to droughts in Africa during the 1980s (Nowak, 2002). These examples highlight the various complexities in the interaction between environmental decisions and societal stakeholders as the nature of response may take years to be felt, the population impacted maybe different from the population which is affected by the environmental decisions and different stakeholders will have different ability to adapt to the decision. Hence it is required to move beyond simple cause-and-consequence to understand these not-so-direct, secondary or tertiary, impacts of environment related decisions.

**Defining ‘indirect’ consequence**

The dictionary definition of ‘indirect’ refers to ‘having intervening factors or persons or influences’, not leading by a straight line or leading through different lines but descending from a common cause. Analyzing society related (stakeholders) consequences through these definitions one can define indirect consequence of environmental legislations as those factors or influences that directly descend/ emerge from the direct consequences, and lead to changes (movement) towards effective/better management of environment and society (socio-cultural aspects). Example of some of the indirect consequences of end of life legislation may include changes in organizational accounting practices or creation of specialized role in the organization structure.

**Defining ‘unintended’ consequence**

The dictionary definition of ‘unintended’ refers to “any activity or influence not done or made or performed with purpose or intent”. The consequence is not part of and has not been planned in the activity or influence in question. The nature of the word suggests “unanticipated effect that could be positive or negative and which leads to making discoveries, by accident and sagacity, of things not in quest of or an effect which could be source of further problems. For example, United States had imposed quotas on imports of steel in order to protect steel companies and steelworkers from lower-priced competition. The quotas did help steel companies. But they also made less of the cheap steel available to U.S. automakers. As a result the automakers had to pay more for steel than their foreign competitors do. So policy that protected one industry from foreign competition makes it
harder for another industry to compete with imports (Norton, 2006). Possible causes of unintended consequences include world’s inherent complexity (leading to either ignorance or incomplete analysis), contrary incentives like in the example of US steel industry, or cognitive or emotional biases, i.e., immediate interests of pressure groups or social values (Merton, 1936).

Therefore for the review we identify unintended consequence as change that may not directly descend from a direct effect, and that creates some hindrance for effective/better management of environment or society (socio-cultural aspects).

**Definition of ‘unforeseen’ consequences**

Unforeseen consequence is defined as “unanticipated and disconcerting lines of development” (Glidden, 2003). First hinted at by Adam Smith (1904) in the Wealth of Nations, while he used the term “invisible hand”, it was used to describe ‘unforeseen’ only once, in the following quotation:“...[B]y directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as in many other cases, led by an invisible hand to promote an end which was no part of his intention. Nor is it always the worse (or good) for the society that it was not part of it.” Within the understanding from Glidden (2003) and Smith (1904), we can define it as “any sudden or unexpected development that has the potential of creating eco-socio-cultural imbalance in the society”. Additionally the magnitude of change would be larger in ‘unforeseen’ than ‘unintended’.

**Theoretical frameworks**

The economic-liberty perspective or economic theory of regulation

Sometimes known as the private interest perspective, it believes that market is the best mechanism for maximising social and economic welfare. It treats political and bureaucratic motives with suspect and highlights the role of interest groups in regulation formation (Wilson, 1980). It believes that in a democratic system governments will establish policies to not only cater to interests of the general public but also to satisfy interests of specific pressure groups and every industry or occupation that has enough power to utilise the state will try to get a favourable regulation in place (Stigler, 1971). Some of the salient features of regulations, from this perspective, are:

Just like elected officials (for votes) and appointed officials (for wealth), industry also works as a pressure group in order to acquire regulations for their benefit.

Often regulations get formed keeping in view the interest group’s perspective rather than of the ones who are to be regulated.

Even if a group has a strong incentive to organise as a pressure group it must still acquire and use influence.

The virtues of this perspective on regulation are that it cuts away the naïve assumptions that government officials (elected or appointed) are selfless, altruistic individuals and that they too respond to awards. Therefore understanding of stakeholder behavior towards a regulation can be predicted by understand possible rewards for that stakeholder.
**The normative-positive perspective or public interest theory**

This perspective looks at regulation as an instrument to overcome or check market failure. Regulations are seen to improve economic efficiency and promote social values by correcting market imperfections like natural monopoly, asymmetric information, using of public good, moral hazard, unnatural transaction costs or creation of externalities (Kearney & Merrill, 1998).

Government intervention in the economy and regulation of environment is required due to presence of externalities like pollution, waste etc, and their presence constitutes an obstacle in optimum resource allocation. The state has several means for potential intervention including economic instruments like taxes, subsidies and other incentives and regulations which have directives and penalties. (Buchanan, 2003; Tanguay et al., 2004).

Some of the market imperfections/failures are:

- Natural monopoly situations (example, power sector) where the monopolist will raise his cost and tariffs as much as possible to take incentive on his efficiency and also maximize his profit.

- Private market activities are seen to create spillovers and externalities which include any cost or benefit not accounted for in the price of a goods or services. A positive externality is when the producer cannot take all the benefits of the activities he has undertaken (example, R&D activity) and a negative externality is when the producer cannot be charged all the costs or producer makes more goods than what is socially beneficial. Both externalities need market regulation to be more efficient.

- Consumption of public goods can create problems of either free riding or excess utilization which needs regulatory intervention.

- People have different/asymmetric information at the time they act making markets inefficient even when there is advantageous transaction that can be made. For example, in a used car sale, the seller knows the actual value of that car but the buyer can only estimate the actual value from the limited information he has about used cars. Because of this asymmetry of information, the buyer is only willing to pay average of the values of the cars he believes are offered for sale. Therefore seller of a car of higher value than the average is at a loss and he may not either offer the car for sale or take lower price.

- Transaction costs are costs associated with making market transactions like reverse logistics, recycling, transportation in the case of EOL legislation. To the extent consumers and producers will incur costs to become informed about the market and also complete transactions, market will not perform efficiently.

- Moral hazard is another imperfection which refers to presence of incentives for individuals to act in ways that incur costs that they do not want to bear. For example, a person with government sponsored medical insurance may not take care preventive care of health unless there are structures to ensure that he takes socially efficient preventive measures.

---

**The pragmatic-administrative perspective**

This perspective takes the presence of governments and markets as the best possible of all available options to the society for its regulation. Instead of dealing with normative and philosophical questions involved, it concentrates on the study of empirical, day to day
problems of regulation as a system of governance. This perspective says that each regulation has a life cycle and within the cycle it goes through similar issues and concerns and governance is about managing the regulation through its various life cycles in the best possible way (Barry, 1980; Bernstein, 1955). An example of this is Bernstein’s (1955) analysis of regulatory life cycles where he argues that all regulations go through a similar life cycle which consists of periods of growth, maturity and decline. He identifies four major periods in the life of any legislation, namely,

- Gestation: lasting about 20 or more years in which a rising distress leads of formation of active groups who demand changes and remedies.
- Youth: where the agency is crusading and aggressive and operates in a conflict prone environment. Typically when the legislation is new, the agency struggles with vague objectives, limited experience and untested areas, while facing opposition for the legislation
- Maturity: the agency undergoes a process of adjustments and attempts to build good relations with other agencies and groups.
- Old Age: the agency develops a set working arrangement with other parties that leads to maintenance of status quo and also establishes the agency’s role.

Theory of unintended consequences or Chaos Theory
Scientists and mathematicians frequently advance the Chaos Theory to explain unanticipated consequences. It states that tiny variations in beginning conditions can trigger mammoth, lasting transformations in the end results. Its basic principle is sometimes referred to as the Butterfly Effect – as in, say, a hurricane being whipped up by ever-increasing wind vectors over a period of days and weeks after being catalyzed by the tiniest imperceptible change in air current when a butterfly somewhere flaps its pretty little wings (Hanchette, 2003).

This theory is based on the understanding that because we live in complex systems that are difficult to understand in totality, our choices can have system-wide implications which we neither intended nor expected. Chaos theorists start with understanding individual choices based on self-interest, but their primary interest is in how these actions affect society as a whole, that is, do these choices lead to chaotic results, or to harmonious ones (Hanchette, 2003). Their concern with unintended consequences of human choice and action leads them to argue that good results do not necessarily come from good intentions, and that good intentions do not necessarily lead to good results (in contrast to popular cultural believe intentions determine results).

Ecological theory of modernization
Ecological theory of modernization has been offered as a possible explanation and understanding of a way to solve environmental problems faced by industrial countries. It suggests that regulation can help to solve environment problem at the same time make industry more competitive. This can be achieved if regulation encourages development of innovative technologies and production techniques (Huber, 1982). It is concerned with the relationship between developments in the industry and the environment with respect to the capacity of modern societies to recognize environment problems and issues. Ecological modernization has been identified as one of the ways in which late modern society is
responding to its increased awareness of and anxiety about ecological risks associated with modernism. This theory is a good base to analyze emergent policy discourses and as a basis from which various policy prescriptions can be brought forward to be analyzed.

## Review

### Plant productivity

<table>
<thead>
<tr>
<th>Authors</th>
<th>Industry</th>
<th>Method</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finnveden et al., 2004, EU</td>
<td>Waste management, Life cycle assessment</td>
<td>The paper studied waste management using life cycle assessment techniques and found a need for creating fresh policies for plastic using manufacturing firms, to support replacement of plastics made from virgin material, which in turn will lead to decreased use of total energy and increased productivity of the firms.</td>
<td></td>
</tr>
<tr>
<td>Knight et al., 2005, USA</td>
<td>Wood products industry, Life Cycle Inventory (LCI) method</td>
<td>The paper compared potential environmental effects in production of wood and steel doors and found that production of steel doors though was saving wood, resulted in higher energy use for its processing, an unintended consequence of trying to save wood.</td>
<td></td>
</tr>
<tr>
<td>Scarpetta et al., 2003, OECD countries</td>
<td>Electrical and Electronic industry, Econometric analysis of secondary, firm level data</td>
<td>The authors studied labour and product markets to analyze the possible role of institutions and regulations on multifactor productivity. They found: stringent regulatory settings in the product market indirectly affected entrepreneurial activity which in turn had a strong negative effect on market access and entry of small and medium size firms. The burden seemed to be greater, the further the industry or firm was from the technology leader. The specific contribution of firm performance and firm dynamics on multifactor productivity of an industry varied across industries and countries and while regulations hindered productivity directly, they had an indirect negative effect on productivity by affecting innovation activity and R&amp;D. Industry level analysis of effect of regulations on various industries shows capital intensity and entrepreneurial activity within firms got affected negatively by stringent regulations.</td>
<td></td>
</tr>
<tr>
<td>Greenstone, 2002, USA</td>
<td>Manufacturing sector, Secondary data analysis from the Census of Manufactures (1970-87)</td>
<td>The paper estimated the impacts of Clean Air Acts, by dividing various US counties into pollutant-specific non-attainment and attainment categories, on measures of industrial activity, obtained from 1.75 million plant observations. The estimates suggest that in the first 15 years in which the regulation was in force (1972–87), non-attainment regions lost significant amount of jobs, capital stock, and output in pollution-intensive firms, compared to attainment regions.</td>
<td></td>
</tr>
<tr>
<td>Boghe, Electronics</td>
<td>EU directives on use of certain</td>
<td>The paper evaluated EU directives on use of certain</td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>Country</td>
<td>Industry/Technical Paper</td>
<td>Summary</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>--------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>2001, UK</td>
<td></td>
<td>Technical paper</td>
<td>materials and found likelihood of less sophisticated electronic equipments like TV and audio, also entering the recycling chain, which in turn will indirect effect multifactor productivity and economics of these equipment manufacturers.</td>
</tr>
<tr>
<td>Berman et al., 1999, USA</td>
<td>Manufacturing industry, econometric analysis of plant level data of oil refineries</td>
<td>The paper examined the effect of environmental regulations on manufacturing plants, using data from 1979 to 1992 and found strategy or gaming aspects to regulations where firms may try to preempt regulators from choosing technologies by introducing new abatement technologies, in order to either reduce uncertainty of future regulations or impose costs on either exiting or potential competitors.</td>
<td></td>
</tr>
<tr>
<td>Dion et al., 1996, Canada</td>
<td>Pulp and paper industry in Canada, Econometric analysis of secondary, plant level, monthly data</td>
<td>While analyzing the determinants of regulators monitoring activities and the factors that explain the decision to inspect or not inspect environmental compliance of producers, authors found that greater effort is allocated to plants which are perceived to be more polluting. Moreover, variables pertaining to local conditions have an impact on monitoring behavior of regulations including local labour market conditions. These results provide support to public interest theory of regulation.</td>
<td></td>
</tr>
<tr>
<td>Jaffe et al., 1995, USA</td>
<td>Manufacturing sector (US dept. of Commerce)</td>
<td>The paper analyzed data/information from Department of commerce to understand competitiveness of US manufacturing firms and finds that interpretation of the laws by a state had an indirect affect on a firm’s decision of plant location.</td>
<td></td>
</tr>
</tbody>
</table>

### Society/Governance and Consumers

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Industry/Conceptual paper</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson et al., 2005, EU</td>
<td>Manufacturing sector, Conceptual, based on literature review</td>
<td>The paper discusses the changing roles, responsibilities and positions of various industrial actors in the process of material flow for recycling. The authors found creation of new actors in the system for fulfilling recycling tasks, namely, suppliers of material for recycling and also emergence of a new, active and involved role of the end consumer, just passive recipient of goods. With increasing recycling activity, some of the preconditions for business models and thinking will have to be questioned. In general, the conceptual thinking regarding the “final” buyer or end consumer, constructed around the idea of a ‘chain’ of actors will have to undergo change towards a circular model. The introduction of recycling not only challenges this linear conceptual thinking but also challenges the business models and concepts based on the liner model. The dual role of the consumer will involve role in the physical flow of material, as a supplier</td>
</tr>
<tr>
<td>Reference</td>
<td>Paper Type</td>
<td>Industry/Sector</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Herold et al., 2005, EU</td>
<td>Technical paper</td>
<td>Furniture retail industry</td>
</tr>
<tr>
<td>Hicks, 2005, China</td>
<td>Discussion paper</td>
<td>Electronic industry</td>
</tr>
<tr>
<td>Cardinali, 2001</td>
<td>Research paper</td>
<td>Based on literature review of regulatory planning models</td>
</tr>
<tr>
<td>Jaffe et al., 1995, USA</td>
<td>Manufacturing sector (from US dept. of Commerce)</td>
<td>Based on literature review</td>
</tr>
<tr>
<td>Meyer, 1995, USA</td>
<td>Manufacturing sector, Econometric analysis of secondary data.</td>
<td>Based on literature review</td>
</tr>
<tr>
<td>O’Riordan, 1992</td>
<td>Manufacturing sector, Research paper based on literature review</td>
<td>Based on literature review</td>
</tr>
<tr>
<td>Study Source</td>
<td>Sector</td>
<td>Methodology</td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>-------------</td>
</tr>
<tr>
<td>Triebswetter et al., 2004, EU</td>
<td>Manufacturing sector, Case studies of three countries</td>
<td>The article examined three case studies to understand negative impacts of environment legislation and found while there was no significant impact on the manufacturing sector, it created positive employment effects together with environmental benefits.</td>
</tr>
<tr>
<td>Huisman et al., 2004 USA</td>
<td>Electronic Industry Primary survey of 75 products</td>
<td>The paper explored outcomes of the eco-efficiency concept associated with end of life legislation in the electronic industry. Other than direct implications like technical constraints, sorting and separation costs, disassembly times and costs, it also found that for substantial environmental gain in relation to financial investments made economies of scale was an essential criterion that had to be fulfilled for the glass and plastic recycling scenarios. This would mean exclusion of small and medium manufacturers of many electronic items unless they expanded their business to a certain level. This requirement of end of life legislation not only unintentionally forces the industry to have only big players but consequently also ends the existence of small and medium firms. Other findings were that for small and medium size housing, the extra cost of plastic recycling are higher than the environmental benefit realized.</td>
</tr>
<tr>
<td>Dept of Trade and Industry, 2003, UK</td>
<td>Automotive Industry, Technical paper</td>
<td>The paper attempted to understand and identify concerns on how existing regulations are affecting competition, in the passenger car segment. While it listed some direct consequences of end of life legislation like changes in market outcomes in terms of prices, quality and service, it also highlighted changes in the nature of competition and advertising from certain product characteristics over another.</td>
</tr>
<tr>
<td>Kim, 2002, EU</td>
<td>Automotive industry, Technical paper on EOL policy (study of related instruments &amp; mechanisms)</td>
<td>The paper explored determinant factors for effective end of life vehicle policy. It found direct factors like extended producer responsibility, importance of monitoring and role of authorities, and indirect factors like cooperation among various actors, namely, dismantlers, producers and monitoring authorities as significant for effectiveness of this regulation.</td>
</tr>
<tr>
<td>Fare et al., 2001, USA</td>
<td>Manufacturing sector, Econometric analysis of industries</td>
<td>A Malmquist-Luenberger productivity index was employed to look at market output and the output of pollution abatement activities of U.S. state manufacturing sectors for 1974-1986. it found that when accounting for change in emissions, average annual productivity growth was 3.6 percent, whereas it was 1.7 percent when emissions were ignored. However, they also found that adjusted productivity growth was higher in states with rapidly growing manufacturing sectors states with slow growing manufacturing sectors. This highlights that while entry of new firms increased productivity as they enter with better technologies, however higher technological costs also create entry barriers for medium firms.</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Industry/Field</td>
<td>Description</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>--------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Wubben, 1999, EU</td>
<td>Chemical industry, case study</td>
<td>The paper detailed the impact (European) environmental legislation on competitiveness of a large chemical company. It argues that environmental legislation has created a rat race among the major players at the (longer-term) cost of smaller companies and innovation.</td>
</tr>
<tr>
<td>Richards, 1997, UK</td>
<td>Electronics industry, Conceptual paper</td>
<td>The paper examined some of the pressures faced by the electronic industry. While it highlights obvious direct effects like need for reevaluation of processes and designs, it also highlights indirect and unintended pressures and changes facing the industry including increased cost of energy to operate environmentally safe processes, initiation of collaborative management practices among competitors for better life cycle management, and generation of new business opportunities in this direction, adding different dimensions to inter firm competition.</td>
</tr>
<tr>
<td>Roarty, 1997, EU</td>
<td>Manufacturing sector, Review paper</td>
<td>The paper examined factors that influence corporate behavior towards the environment and found that one of the consequences of setting new environment rules is creating fresh financial structures (taxes and subsidies) for the industry such that market values reflect society’s preferences. This result lends weight to market failure /imperfections theory of legislations. Further, it highlights the new role of consumers and investors as not only recipients of goods and services but as actively involved in defining what they want and how they want it.</td>
</tr>
<tr>
<td>Zhuang et al., 1997, UK</td>
<td>Chemical industry, Firm level (n=203) primary survey</td>
<td>The paper surveyed 203 chemical firms to understand the legislation–firm behavior incongruity in following environmental regulations and found various reasons for the incongruity including absence of governmental incentives and subsidies. This result highlighted the need for governmental intervention to motivate firms, through introduction of subsidies and taxes in an otherwise market regulated industry.</td>
</tr>
<tr>
<td>Barnes, 1994 UK</td>
<td>All industries, General review</td>
<td>The paper examined how regulations affect businesses of all sizes. While it listed some short term costs for all firms it found specific concerns of small enterprises where costs associated with environmental procedures itself like paperwork and legal fee of consultants and registration, unintentionally, made their prime activity economically feasible. Only large firms were able to bear the cost of environmental procedures.</td>
</tr>
<tr>
<td>Venables, 2005, EU</td>
<td>Electrical and electronic industry, Conceptual paper</td>
<td>The paper looked at the dilemmas facing manufacturers in the electronic industry and found need for creation of indirect incentives for manufacturers to innovate and in that provide opportunities for business advantage, through supportive legislations.</td>
</tr>
<tr>
<td>Karmierczak et al., 2004, EU</td>
<td>Automotive industry, literature search</td>
<td>Objective of the paper was to present the ergonomics of disassembly production systems and the author found nature of communication channels between dismantlers and design engineers as indirectly influencing the production systems.</td>
</tr>
<tr>
<td><strong>International Trade</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Togeiro et al., 2004, Brazil</strong></td>
<td>Food industry, review of trade/ environmental regulations</td>
<td>The paper analysed global environmental regulatory constraints on Brazilian exports and found that environment regulations indirectly affected productive chains in Brazil, especially of small and medium size companies, and agriculture as agriculture based products are exported to developed countries from Brazil.</td>
</tr>
<tr>
<td><strong>Van Beers et al., 2003, EU</strong></td>
<td>Econometric analysis from several industries</td>
<td>The study looked at environmental regulation impacts on international trade The aggregate results support Tobey's findings, namely that no significant impact on international trade is caused by stricter national environmental policies.</td>
</tr>
<tr>
<td><strong>Esty, 2001, USA</strong></td>
<td>Multiple Industries, Literature Review</td>
<td>The paper studied various environmental concerns in trade like food safety requirements, waste management and disposal rules, and recycling regulations and suggested that expanded trade liberalisation could worsen environmental conditions, like creating international level free riders who will be difficult to track, creating imperfect trade transactions, and imposing unwanted ethical and psychological constraints.</td>
</tr>
<tr>
<td><strong>Fernie et al., 2001, UK</strong></td>
<td>Food Industry, Literature review</td>
<td>The paper compared the impact of environmental regulations on various food retailers. It concluded that each country interprets these regulations in its own way and these regulations are seen to create trade barriers. For example, as drinking cans are banned in Denmark, export of beer to Denmark is economically feasible for only those producers which can afford alternative packaging.</td>
</tr>
<tr>
<td><strong>Low and Williams, 1998, UK.</strong></td>
<td>Electronic industry, Technical paper,</td>
<td>The paper examined consequence of end of life legislation on suppliers in the UK electronic industry. It found that other than direct effect of changes in product design, end of life legislation would impact the commercial position of more than 45% suppliers from the Far Eastern countries as the legislation would force them to take products back, remanufacture in Europe or accommodate the negative effects of transportation making the trade economically unfeasible. This would unintentionally impacting trade between Far East and UK.</td>
</tr>
<tr>
<td><strong>Braithwaite, 1995, Canada</strong></td>
<td>Electronic industry, Case Study</td>
<td>The paper studied an entrepreneurial activity undertaken by two entrepreneurs and found new trade opportunities in secondary electronic goods to countries which have less stringent environment regulations.</td>
</tr>
<tr>
<td><strong>Prendergast G. P., 1995, UK</strong></td>
<td>Consumer goods industry, survey of UK marketing executives</td>
<td>The paper surveyed marketing executives from consumer goods firms to understand the relationship between packaging, environment and logistics. Other than substantial cost increase as a direct effect of environment directives, the author also found respondents concerned that logistical complexity under the environment directive would create trade barriers.</td>
</tr>
<tr>
<td><strong>Candice et al., 1994, OECD</strong></td>
<td>Life cycle assessment approach</td>
<td>The paper studied environmental life cycle / trade and found that policies were sometimes unintentionally discriminating against imports, particularly from developing countries.</td>
</tr>
<tr>
<td>Author(s) and Year</td>
<td>Industry, Study/Field</td>
<td>Summary</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Grossman et al., 1993, USA</td>
<td>Export-import with Mexico, Secondary data across industries</td>
<td>The paper studied impact of US pollution related regulations on export-import between Mexico and found that pollution costs related regulations did not have any direct or indirect affect on trade between the two countries.</td>
</tr>
</tbody>
</table>

### Organizational systems and processes

<table>
<thead>
<tr>
<th>Author(s) and Year</th>
<th>Industry, Study/Field</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cerin, 2003, US and EU</td>
<td>Telecom industry, Case Study of Ericsson</td>
<td>The paper studied the case of Ericsson to understand linkage between environment and corporate economic benefits. While studying the direct effect of life cycle assessments (LCAs) on development of more eco-efficient designs, the author found effectiveness of eco-efficient designs requires change in corporate planning systems and formal organisation behavior towards integrating environment concerns at all levels.</td>
</tr>
<tr>
<td>Dept of Trade and Industry, 2003, UK</td>
<td>Automotive Industry, Technical paper</td>
<td>The paper identified concerns on effect of existing regulations on competition in the passenger car segment. While it listed some direct consequences of end of life legislation it also highlighted need for new systems and processes within the organisation for effectiveness of activities associated with environmental compliance.</td>
</tr>
<tr>
<td>Simon et al., 2000, UK</td>
<td>Electrical and electronic industry, DEEDS project</td>
<td>The paper looked at integration of environment issues into product design. While the authors suggest a four stage framework for eco-design practice, they also highlight organisational requirement of more complex communication structures and new roles (other than marketing) which will champion the environmentally friendly designs where the responsibility would also include acting as source of expertise and as a channel of communication when knowledge transfer is required.</td>
</tr>
<tr>
<td>Miller et al., 1998, USA</td>
<td>Manufacturing sector, survey of 200 firms</td>
<td>The paper studied effects of compliance on business operations and found that compliance led to changes in existing organizational practices and work methods.</td>
</tr>
<tr>
<td>Prendergast , 1995, UK</td>
<td>Consumer goods industry, Primary survey of UK marketing executives</td>
<td>The paper surveyed marketing executives from consumer goods firms to understand the relationship between packaging, environment and logistics. Other than substantial cost increase as a direct effect of environment directives, the author also found increased importance of communication between recyclers and end users and hence need for development of formal communication structures between them. He also highlights emergence of new industry structures in Germany and US, where competitors within an industry develop joint ownership of recycling systems, at various levels, and each firm pays a fee according to their recycled volume.</td>
</tr>
<tr>
<td>Coulson et al., 1995, UK</td>
<td>Finance Industry, Technical paper</td>
<td>The paper looked at the implications of legislative and market pressures on finance lenders and equity investors. It found that lenders may suffer indirectly because of the effect of environmental liabilities on borrower’s solvency or the security of borrower’s loan. It gives the example of Maryland Bank in USA which had a huge loan go into default at the end</td>
</tr>
</tbody>
</table>
of 1991 due to borrowers’ in ability to meet environmental requirements. Some of the implications for lending institutions include increase in costs to carry out rigorous risk assessments, creating structures to encourage borrowers to turn green like ethical savings accounts, discounts etc., and putting investor money in companies directly involved in environmental areas like waste management.

REFERENCES


Economic and Social Research Council, (2004). The UK's productivity gap: what research tells us and what we need to find out. Swindon, UK.


An Exploratory Investigation into the Impact of Legislation and Regulation on Industrial Organizations: Tentative Speculations from reviewing ‘The Engineer’

Naomi Brookes

Aim and Introduction

The aim of this investigation was to generate a tentative perspective on the effects of legislation and regulation on industrial organizations that could be used to inform further research investigations. Using a conventional ‘systematic review’ approach to examine abstract databases (i.e. ABI Inform and INSPEC) had proven lengthy and problematic. Therefore, in order to gain a ‘first-pass’ viewpoint, a cruder but quicker method was employed.

Instead of searching across many publications using abstract databases, a search was conducted in a single publication, ‘The Engineer’. ‘The Engineer’ is a suitable journal to provide an overview of effects on industrial organizations as its focus of concern covers the spectrum of organizations in which professional engineers participate. (If the exploratory investigation was to be conducted with more rigor the circulation figures by industrial sector would be needed to justify this assertion.)

The terms ‘unintended’, ‘indirect’ and ‘UK productivity’ were not operationalised and employed in this exploratory investigation. The concept of ‘unintended’ effects of legislation is difficult to define. (For example, is the rise of alternative car engine designs an intended or unintended effect of emission regulation?) UK productivity is an equally difficult concept to make concrete. (For example, should productivity be used in a precise sense e.g. output per employee or should effects that lead to increased costs also be included?) for these reasons any link between legislation and regulation and an effect in one or more industrial sectors was noted. However for the purposes of the wider project, the concepts of ‘unintended’, ‘indirect’ and ‘productivity’ need to be revisited and defined.

Method

The Engineer Online www.theengineer.co.uk (searching this site actually includes not only the print content of The Engineer but also the web-content of The Engineer Online and The Engineer’s sister publication Process Engineering,) was searched using the ‘advanced’ search feature available using the term ‘legislation OR regulation.’ and date limited from 1st January 2004 to 31st January 2006. The identified abstracts were then reviewed manually to eliminate those that did not directly relate legislation or regulation to a specific effect in one or more industrial organizations. These were then categorized in inductively generated categories from the review of the remaining abstracts. The distribution of abstracts across these categories was used to generate tentative speculation about the impact of legislation and regulation on industrial organizations.
Results

The emergent categorization schema is given in Table 1. The distribution of abstracts from the publication within that categorization schema is shown in Figure 1 a) – d).

<table>
<thead>
<tr>
<th>SECTOR AFFECTED</th>
<th>TYPE OF LEGISLATION</th>
<th>EFFECT OF OR LEGISLATION/REGULATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pharmachem &amp; Process</td>
<td>Beneficial</td>
<td>Environmental – General</td>
</tr>
<tr>
<td>Automotive</td>
<td>Inimical</td>
<td>Environmental – Emissions</td>
</tr>
<tr>
<td>Energy &amp; Utilities</td>
<td>Leading to the introduction of a specific product or technology</td>
<td>Environmental – End-of-life</td>
</tr>
<tr>
<td>Environmental</td>
<td>Leading to the generation of a new market</td>
<td>Environmental – IPPC</td>
</tr>
<tr>
<td>Electrical &amp; Electronic</td>
<td>Beneficial due to de-regulation or de-legislation</td>
<td>Health and Safety</td>
</tr>
<tr>
<td>Others</td>
<td></td>
<td>Others</td>
</tr>
<tr>
<td>Trans-Sector</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: The Emergent Categorization Schema

![Figure 1a](image1a.png)

**Figure 1a:** The Distribution of Abstracts Categorized by Sector

![Figure 1b](image1b.png)

**Figure 1b:** The Distribution of Abstracts Categorized by Legislation/Regulation Effect
Conclusions and Further Work

A number of very tentative conclusions can be reached regarding the effects of regulation and legislation on industrial organizations within the UK:

- Automotive, Environmental Industries and Phrmachem and Process Industries appear to be Most Affected. Industry sectors that have been most affected by legislation and regulation, based on the number of abstracts in each categorization, appear to be ‘Automotive,’ ‘Environmental’ and ‘Pharmachem and Process.’ (This, of course, assumes that the distribution of all abstracts across all sectors in the publication in question is uniform. Before more certain conclusions could be drawn, this would have to be ascertained.)

- Regulation and Legislation Appears to Stimulate Innovation. Reported effects of legislation and regulation do not reflect a situation where their introduction necessarily results in problems to industrial organizations. Instead, the introduction of legislation and regulation appears to act as a spur to innovation. A number of examples have been identified where new technologies, products and markets have been developed in direct response to regulation or legislation. Where inimical effects have been reported these tend to be on a trans-sectoral basis.

- Environmental Legislation and Regulation Appears to be Most Impactive. Assuming that it is safe to judge relative impact of legislation/regulation type by the number of abstracts in a particular category, environmental legislation and regulation has by far the greatest impact on industrial organizations. The methodological mechanism behind this assertion does need examination however. For example, if ‘Personnel Today’ had been chosen as a basis from which to form a perspective on the impact of legislation then employment legislation may be the most frequently quoted.

A number of ‘micro’ case-studies to explore the effects of legislation and regulation are suggested by this investigation. These include the following:

- the effects of the return to steel fuel tanks in cars (rather than zinc alloys) following the introduction of end-of-life regulations
- the impact of ‘end-of-life’ regulations on fasteners
- the interaction between environmental legislation and the growth of the environmental catalyst market
Using the approach of searching within a sample population of specific publications appears to be a feasible mechanism for essaying a thematic representation of the effects of regulation and legislation on UK industry. A potential methodology in presented in Figure 2.

For each publication:
- Identify a sample population of 'trade' magazines and journals
- Search within publication for 'legislation OR regulation' in defined time period
- Manually review abstracts to those that relate legislation/regulation to its effect
- Inductively categorise and analyse abstracts (c.f. grounded theoretic approach)

Combine analysis with existing analyses to:
- identify overall themes
- identify potential case-study areas

Figure 2: Research Methodology for Thematic Analysis of Effects of Regulation and Legislation on UK Industry
AIR TRAVEL DE-REGULATION AND LOCAL ECONOMIES IN EUROPE: REFLECTIONS ON THE UNINTENDED CONSEQUENCES FOR PRODUCTIVITY

Vlado Balaz and Allan Williams

Introduction
The dominant feature of the air travel sector in recent decades has been a series of deregulatory initiatives. These were meant to increase competition, resulting in lower prices and welfare gains for air travellers. These changes were manifested in both the economic behaviour of existing scheduled airlines and the spectacular growth of new market entrants, the low cost carriers. The result has been a striking increase in the scale of air travel. According to IATA (International Air Transport Association), the demand for air travel in Europe increased three-fold between 1980 and 2000, and is set to double by 2020 (http://www.iata.org/pressroom/industry_facts/stats/2003-04-10-01.htm).

But the changes in air travel are not only quantitative but also qualitative in the sense of creating new geographies of air travel and accessibility, evident in differential growth of airports, routes, passenger flows, and hub versus spoke connections. This paper considers how these changes have had unintended consequences for productivity levels in particular local economies.

The very beginning of low cost air travel (LCA) is usually traced back to 1967, when the Southwest Airline was founded. Flight operations began in 1971 with flights within Texas. Its business design included short-trip service, low prices, high point-to-point frequencies, high punctuality, no frills and a brand image that “flying is fun”. From the 1980s, LCAs became a significant presence in European air travel. LCAs currently account for around 24% of the scheduled intra-European air traffic and the market share of LCAs continues to grow strongly. The United Kingdom has the most developed market for low fares services, with the low fare share of scheduled traffic at the airports approaching 50% in 2004 (UKCAA 2004). This success has been replicated, more recently, in several other national markets for air travel within Europe.

LCA achieved intended increases in productivity via a number of organizational and institutional innovations. According to Francis et al (2005: 84):

‘The core characteristics common to the majority of low cost airlines are: high aircraft utilization, internet booking, use of secondary airports, minimum cabin crew, lower wage scales, lower rates of unionization among employees, one class of seating thus allowing more seats per aircraft than traditional airlines (who offer alternative seat pitches for different classes of travel), short ‘on the ground’ turn around times, no cargo carried to slow down turn around times, a simple fare structure and pricing strategy, e-ticketing, no seat allocation, passengers having to pay for food and drink, flexible working terms and conditions for employees relative to traditional airlines, point to point services and no connections offered’

But there have also been a number of unintended productivity changes, beyond the airline sector itself. Here we consider the consequences in terms of productivity levels outside of the airline sector. The key to understanding this is recognition of the space and time specificity
of the changes. Docherty (2004: 341) sums up the relationship between air travel links and regional competitiveness:

‘Much recent research generally supports Porter’s position. After economic diversity, high quality internal and external connectivity has been noted by several studies to be the most important explanation of regional economic competitiveness. The most successful regions have class-leading transport and ICT infrastructure to move goods, services, information and people securely, quickly and efficiently. Particularly for knowledge intensive industries, the presence of direct international air links to key global centre’s of innovation is regarded as critical. Despite the new opportunities presented by ICT, face-to-face communication remains fundamentally important to the sharing of ideas, and to the development of trust on which all business relationships depend. Spinning out from this connectedness was an outward looking approach, with many of the most successful regions investing significant time and effort in developing links with other cities to share knowledge, open new markets, and generally to win friends and influence people. Internal connectivity was also deemed crucial, with congestion becoming an increasingly influential constraint to business investment and quality of life.’

Their approach is essentially ‘black box’ in respect of air travel itself, which is treated as uniform or undifferentiated. We argue that, when considering local economic consequences, ‘air travel’ or ‘volume of air travel’ is a chaotic conceptualization (Sayer 1992). There have been strongly differentiated changes in the types of passengers carried: between existing scheduled carriers and LCAs, amongst LCAs, and amongst airports. In particular, and in relation to unintended productivity impacts, we can distinguish between three main types of travelers: labour migrants, tourists, and business travelers. The changes in accessibility and mobility realized by each of these groups has had differentiated productivity impacts for local economies.

In part, the unintended productivity consequences of the growth of budget airlines are related to the highly uneven economic structure of space. In the simplest terms, the economy of a local space is constituted of a unique combination of basic production factors, namely labour, capital and ‘other factors of production’, (amongst which, we are chiefly interested here in knowledge). These production factors are neither immobile nor interchangeable, and economic geography increasingly emphasizes the notion of ‘economies of flows’ (Hudson 2004). Changes in air travel potentially can have substantial impacts on changes in the volume, structure and effectiveness of production factors in particular spaces, with important consequences for productivity. However, as will be seen from the following review of air travel de-regulation, most if not all of these productivity consequences – beyond the air travel sector itself – were unintended.

The specific objectives of this paper are to review

- the unintended impacts of air travel deregulation on flows of trade, investment, labour and knowledge
- how these impact on productivity levels in local economies
- and thereby contribute to changes in uneven regional development
Conceptual framework

Competition and regulation
Existing research suggests there are links between deregulation and productivity, but regulation is necessarily complex.

- Regulation is multi-level and these levels are ‘folded in’ on each other (Amin 2002). Hence the impacts of deregulation on productivity at one level will be mediated by regulatory changes or lack of changes at other levels. Regulation occurs at many different levels eg EU and national levels measures to deregulate air travel may be significantly mediated by local planning controls.
- It is not sufficient to de-regulated rights of carriage – you also have to deregulate baggage handling and other services, to realise even the intended consequences of deregulation, i.e. there is need for a series of directly linked de-regulations.
- There is also a need to look at other, apparently unconnected arenas of regulation, for example on the environment (take off and landing noise levels), employment laws etc.

All of this means that ‘geography matters’ (Massey 1984): regulation or deregulation do not occur in a vacuum but instead the impacts, whether intended or unintended, are place and time specific. For example, the effects of the introduction of LCAs in the 1990s appear to have been far less than of the subsequent expansion in the 2000s. Or the impacts of LCAs on productivity in the Exeter region compared to Glasgow, or in Bratislava compared to Stanstead, are very different.

Local ‘economies of flows’
There are competing conceptualizations of the relationships between flows and spaces. This paper is influenced by Massey’s (1994: 154) view that places are a constituted of local and more spatially stretched relationships, that is that they are ‘articulated moments in networks of social relations and understandings’. This means that social and economic relationships are – at least temporarily – locked into particular places (Allen et al, 1998). This locking into place of particular flows (of human capital, trade, knowledge and financial capital) is constantly subject to reformulation. One of the sources of these changes is air travel deregulation. While probably not the major driver of changes in the flows of production factors in most local economies, it can have significant impacts in particular places and at particular times.

The main flows subject to being reconstituted due to air travel deregulation are:
- Trade, particularly tourism, but also just-in-time deliveries (though note that many LCAs do not carry cargo as a way of speeding up turn around, and intensifying the use of their capital assets).
- Migration – both the volume of flows, and their changing constitution, including more emphasis on circulation and temporary migration.
- The links between LCA and capital flows are complex. However, where a link can be established between LCA growth at particular airports, and inward investment, then the literature on ownership of capital suggests that inward investing firms will have knowledge advantages and higher productivity.
- Depending on the importance of tacit as opposed to codified knowledge (Polanyi 1966), and the role of face-to-face contacts in facilitating transfers of such knowledge (Williams 2006), LCAs can influence the acquisition and transfer of knowledge across regional/national boundaries, either intra- or inter-company.
Productivity

Productivity is of course a measure of the relationship between inputs and outputs, whether for individual firms, or for aggregates such as local, regional or national economies. It is notable that the Porter and Ketels 2003: 11) review highlights the following as important 'intermediate indicators of many other microeconomic attributes of an economy': capital intensity, labour force skills and total factor productivity.

The argument advanced in this paper is that although air travel re-regulation was driven by the aim of increasing competition in order to restructure that sector, bringing about increased competition and productivity, lowering prices and increasing the volume and geographical spread of air travel, it also had unintended consequences for productivity levels in other sectors of the economy. The key argument is that changes in the flows of capital, labour, knowledge/trade, as set out in the previous section, have had unintended consequences for productivity because of the way they have impacted, respectively, on capital intensity, labour force skills, and total factor productivity.

In addition, the paper emphasizes that the unintended productivity effects have been time and place specific. The main EU directives have had nationally and locally variable unintended impacts on productivity because of differences in how economies (in the widest sense, including institutions) are constituted and regulated.

It should also be noted that the resulting unintended changes in productivity levels in local economies will have feedback effects in the form of further unintended productivity consequences for particular carriers. However, this lies outside the scope of this study.

History of air travel deregulation and liberalization

Economic agents (e.g. airline companies) had few incentives to innovate in the rigid EU air market in 1980s and, instead, adapted to operating within a framework of generous public subsidizes and protective policies relating to national air transport infrastructure (e.g. landing and take-off slots). Some European companies, however, learned from the US experience of re-regulation. Applying 'innovation by imitation', they were able to anticipate ongoing changes on the global air market, and lobby for changes in European regulation.

The air transport sector in the European Union was liberalized in three successive stages. The 'first package' of measures, adopted in December 1987, started to relax the established rules. For example, it limited the right of governments to object to the introduction of new fares. Some flexibility was allowed to enable airlines in two countries, that had signed a bilateral agreement, to share seating capacity. Previously, absolute parity had been the rule. In June 1990 a 'second package' of measures opened up the market further, allowing greater flexibility over the setting of air fares and capacity-sharing. Moreover, the new provisions extended the right to the fifth freedom and opened up the third and fourth freedoms to all Community carriers in general. These measures, which were initially limited to passengers, were extended to freight in December 1990. The last stage of the liberalization of air transport in the EU was subject to the "third package" measures adopted in July 1992, and applied from January 1993. This package was far more radical than the previous ones. Effective from 1 April 1997, all EU carriers have had open access to virtually all routes within the EU (freedom to provide cabotage). Liberalization of air travel was intended to ensure that 'air fares should normally be determined freely by market forces' (Council Regulation (EEC) No 2409/92). Further liberalization initiatives were aimed at the scarcity
and cost of infrastructure, which the European Commission had identified as a major cause of the high costs incurred by European air travelers. As of January 1999, access to the ground handling market was liberalized for third parties at Community airports, as provided for by Directive 96/67. This measure helped to reduce operating costs and improve the quality of service for airport users. The fragmentation of the air traffic management systems was addressed through the ‘Open European Sky’ regulations (Regulation (EC) No 549-552/2004. Non-discriminatory and transparent use of computerized reservations systems was introduced by Council Regulation (EEC) No 2299/89.

**The impacts of liberalization**

**Intended**

Papatheodorou (2002: 384) summarized the reasoning behind air travel de-regulation: In addition to the dissatisfaction with regulation, liberalization was partly justified on the presumption of competition working in thick markets and the significance of contestability …. on thin routes. Under deregulation a multitude of new entrants in popular markets would induce carriers to actively compete. Although the possibility of an initial instability and shakeout was recognized, the number of surviving carriers was regarded as sufficient for competition and for realization by the airlines of their interdependence. Self-defeating and mutually destructive price wars would be avoided. Similarly, some deregulation advocates believed that services would improve in peripheral routes despite the non-sustainability of active competition. The efficiency of an airline monopoly (or powerful oligopoly) relied on the threat of hit-and-run entry by non-incumbents which accumulate if super-normal profits pricing emerged. The presumed low entry and exit costs in conjunction with the inflexibility of the incumbent’s fare structure could facilitate such a speculative strategy by a potential entrant. To preempt this practice the existing carrier would limit prices to a level consistent with normal profits.

Air travel market liberalization has brought about significant changes relating to the growth of LCA. Most of the traditional scheduled airlines were state-owned and generally offered relatively high-cost, low-frequency services, especially to less advantaged locations, often combined with poor connections across capital city airports. The introduction of LCA brought consumers substantial benefits from competition. These included – above all - higher frequencies, cheaper fares and increased connections. By 2004 around 60 new entrant and charter/regional airlines had applied the low fares model to varying degrees in Europe, offering consumers low cost services to many, often previously unconnected, destinations. The average fare of €41 for Ryanair and €62 for EasyJet compared with €200 for Lufthansa and Air France, and €268 for British Airways in 2004 are symptomatic of the increase in competition, although such comparisons have to be approached cautiously. There were also most important changes in ground handling and catering, notably the emergence of the practice of subcontracting main base activities by major airlines, as well as their extensive use by smaller airlines, especially new entrant low-cost airlines (EC 1999).

There are some similarities between air travel deregulation/ liberalization in the USA and EU. The driving force behind the 1978 Airline Deregulation Act in the USA was the perception that regulation by the Civil Aeronautics Board (CAB) had resulted in reduced competition and higher air fares. Although there were nearly three years of congressional hearings preceding deregulation, it is notable that most of the evidence was anecdotal – symptomatic of the lack of hard evidence. US experience has been replicated by the EU. Competition by
the LCA forced the traditional air carriers to apply some features of a ‘lean production’ model: for example, lowering fares, reducing ‘frill’ services on flights, and increased flight booking on the internet (Goetz and Graham, 2004).

While the case for de- or re-regulation was expressed in terms of competition, productivity changes were implicit, and sometimes explicit, in the discussions that surrounded regulatory changes. In practice, re-regulation has had major and largely intended consequences in increasing productivity in the air travel sector – effected both through new entrants having higher productivity than firms departing the industry, and through reorganization within existing companies, the so-called ‘legacy carriers’.

Of course, the anticipated competition effects were not always realized. First, there was a tendency to reduce competition at hub airports (Frenken et al 2004: 233). There are assumed to be associated differences in the consequences for productivity, although evidence is lacking on this point. In addition, there is also evidence that some routes can only support a single carrier, with the resultant emergence of monopolies (Papatheodorou 2002). The consequences for prices and for productivity remain largely a matter of speculation.

**Unintended consequences**

Many of the earlier discourses tended to be negative. Graham (1998: 90) argued that ‘benefits from competition are unevenly distributed’ among European regions. Although liberalization was realizing consumer benefits, the impact was constrained by regional development disparities, which were the key to determining geographical demand for air transport. This argument was partly mediated by the subsequent expansion of LCA to less-favoured regions, including new Member States of the EU. While the ‘hub’ airports had mostly benefited from air travel deregulation in 1990s, competition decreased in these airports in the 2000s, in some respects (Frenken et al 2004). This trend was related, in part, to increasing numbers of regional airports, and flights from these airports, many of which were operated by the LCA.

Goetz and Graham (2004: 275) argue that globalization and liberalization strategies have rationalized the airline industry into a more efficient operation that enhances its long-term sustainability, but in the same time ‘the tactics driven by globalization and liberalization have, at best, unpredictable consequences in terms of sustainability’, as they have resulted in excessive air traffic growth and wasteful competition. Some positive impacts of the air travel deregulation on tourism development were are discussed by Papatheodorou (2002), as noted earlier. This debate about sustainability has implications for how we understand the impact on productivity. First, whether a relatively narrow and conventional approach is taken to outputs, or whether wider social and environmental impacts should also be taken into account. And, secondly, in deciding the appropriate time frame for measuring such impacts.

There have also been unintended consequences for employment within the industry. Several studies have tried to analyze the impacts of air liberalization on job generation. Contrary to expectations that cost rationalization would decrease job numbers in the air industry, the direct impacts have been positive. The overall number of employees in civil aviation increased from 435,400 to 489,700 following restructuring by the airlines, 1988-1996 (EC 1999). According to the Civil Aviation Authority (UK CAA 2004), the evidence from the UK is that liberalization has facilitated the growth of the aviation market and has boosted employment in that sector in the UK. The numbers of employees rose by 15% in British Airways, and by 38% in other airlines, between 1991 and 2001 (UK Civil Aviation Authority 2004). And SkyEurope, for example, estimated that 1 million passengers have generated some
4350 jobs in the Slovak economy, directly or indirectly, 1 SkyEurope: SkyEurope Airlines impact on aviation and Slovak economy, presentation by SkyEurope, Bratislava. Given the increased ‘outputs’ in this period, whether measured in terms of flights or turnover, there have been massive increases in productivity, measured in terms of labour productivity. A similar picture emerges when examining the US domestic industry following deregulation.

The indirect and induced increases in employment have been no less important. Hakfoort at al (2004), for example, found that the combined indirect and induced job multiplier was around 2.0 in case of the Amsterdam airport. It is precisely the productivity impacts which lie behind these and other changes, beyond the airline sectors, that are of central interest to this paper.

**Post rationalization of unintended consequences**

It is often difficult to differentiate between the intended and unintended consequences of de- or re-regulation, because of the prevalence of post-rationalization in many of the attendant discourses. For example, the European Union’s Committee of the Regions (Committee of the Regions, Own-initiative opinion on Low-cost airlines and territorial development, 17 June 2004), recently observed that:

“[t]he availability of regional air services, and in particular low-cost air services, operating from regional airports improves access to the global economy. This, coupled with the lower labour costs and facilities costs associated with the more remote regions, can encourage the business community to locate new economic investment within the region. Existing businesses in the region could develop their market share by being able to reach other parts of the Member State, the EU and the rest of the world”

The Committee concluded that the overall economic impact that can be realised through the development of regional air services operating out of regional airports should be considered under the following four headings:

- **Direct Impact**: employment and income that is wholly or largely related to the operation of the airport;
- **Indirect Impact**: employment and income generated in the economy of the study area in the chain of suppliers of goods and services;
- **Induced Impact**: employment and income generated in the economy of the region by the increased spending of the incomes of these direct and indirect employees;
- **Catalytic Impact**: employers and income generated in the regional economy by the wider role of the airport in improving the productivity of business and in attracting economic activities, such as inward investment and inbound tourism.

This view from the Committee of the Regions either explicitly or implicitly recognizes the impact of de- or re-regulation on productivity levels at the local or regional level. It is, however, a post-rationalization and such consequences were not prominent in the discourses surrounding the enactment of de-regulation. Rather, as emphasized earlier, this was couched largely in terms of competition without reference to the contingencies of time and place.

**Unintended impacts of LCAs on productivity**

Following the classic three fold divisions adopted in productivity studies, the UK’s productivity gap – especially with the USA - can be reduced to three key elements:

- Lower productivity of labour, because of generally lower labour skills
• Lower capital intensity
• Lower total factor productivity (TFP)

The following discussion sets out briefly how changes in flows generated by LCA potentially have had unintended consequences for productivity in relation to these three elements. In the absence of any detailed research on this topic, the discussion is essentially at the conceptual level.

**Labour markets**

Mass migration, of course, pre-dated airline deregulation. However, there have been shifts in the nature of migration in recent decades (King 2002). Longer term migration flows have been replaced by more temporary migration, sequential migration, and cycles of migration. There has also been the growth of various forms of long-distance commuting, involving regular return travel to home regions.

LCAs can impact on labour markets in several different ways, but mainly through reducing travel costs and increasing accessibility. Effectively, it can reduce the transaction costs of international labour migration. The outcome may be increased levels of migration if any of the following conditions apply:

- the costs of travel are a significant barrier to air travel;
- the frequency and convenience (mainly accessibility) of air travel are a significant barrier to air travel; and linked to this
- the availability of cheaper, more frequent or more accessible air travel makes new forms of mobility (punctuated by more frequent return visits) possible to those who would not have been able or willing to engage in more conventional longer-stay migration.

The impacts on productivity are difficult to predict on an a priori basis. The productivity impacts depend on: a) effects on wages and costs, b) filling particular or generalized labour markets shortages, and c) raising or lowering aggregate skill levels in the destination regional economy. There is considerable research on these issues, relating to human capital theories, mainly at the national level (for example, Dustmann, C., Fabbri, F., Preston, I., Wadsworth, J. (2003), Labour market performance of immigrants in the UK labour market, London: Home Office, Online Report 05/03; Dustmann, C., Fabbri, F., Preston, I., Wadsworth, J. (2003), The local labour market effects of immigration in the UK, London: Home Office, Home Office Online Report 06/03). However, there is still very little evidence about the impacts on productivity at the national, let alone the local/regional level. In part this is because the transfer of skills via migration is not only a matter of changes in aggregate levels of human capital in the destination, but also about the social recognition of these skills, and whether migrants have sufficient en-cultured and embedded knowledge to maximize their embodied and embrained knowledge (Williams 2006). This relates to the debates about brain waste versus brain gain (as summarized in International Labour Office (2002), Skilled Labour Migration from Developing Countries, Geneva: ILO, International Migration Programme, International Migration Papers 55 (A M Findlay and E Stewart)

Most of the econometric evidence for the UK suggests there is a time lag before workers can maximize the return to their human capital, due to the need to learn about local practices and institutions, or to acquire language competency (Dustman 1994). There are therefore likely to be similar time lags before the full consequences for productivity are realized.
**Business travel and tacit knowledge**

There are many competing theories of knowledge transfer, including those which emphasize localized learning (Maskell and Malmberg 1999) and those based on non-localized networks, for example communities of practice (Wenger 1998). In reality, most firms probably draw on a range of different networks, at different scales ranging from the local to the international. In this context, the key question is whether proximity is critical for building up the trust that is essential for effective knowledge transfers. Amin (2002: 393-4) argues that physical proximity and localized face-to-face contacts are not essential for trust-based relationships. Instead, intimacy may be achieved, and trust may be fostered, through frequent and regular contacts enabled by the distanced networks of communication and travel as well as the unbroken interplay between face-to-face and telemediated contacts. In contrast, Allen (2000: 28) stresses the importance of human mobility to effect localized networking: ‘the translation of ideas and practices …. (is) likely to involve people moving to and through local contexts, to which they bring their own blend of tacit and codified knowledge’s’. There is surprisingly little empirical evidence as to the relative importance of proximity-based versus distanced knowledge transfers, but in practice these are likely to be complex and highly variable at the level of the individual, the firm and the local/regional economy.

LCAs potentially can impact on the frequency of face-to-face contacts, reshaping knowledge-sharing networks, and the efficiency of knowledge transfer. This may be via increased intra- or inter-firm mobility, attendance at conferences or exhibitions, as a result of lower cost, more frequent and new air connections. Whether this is significant depends on a) the additionality and substitution (with other transport forms for the same trip), b) the importance of face-to-face contacts, and c) the company strategy for disseminating and applying knowledge. However, where these assumptions are met, air travel can be linked to productivity levels. Knowledge is usually recognized in quantitative analyses of productivity as part of total factor productivity. In summary, changes in air travel have changed the geography of business travel, and potentially, therefore, of some channels of tacit knowledge transfer, and hence of total factor productivity.

**Foreign investment, business organization and business travel**

Arguably, both direct and indirect investment are induced by LCA. Our interest here is in the indirect consequences beyond the airline sector. The main effects are:

- LCAs may widen the search spaces of potential investors, who seek minimum levels of accessibility between their different branches.
- The instigation of LCA services may change the image of a local or regional economy, as indicated by the EU’s Committee of the Regions, and this may also attract foreign investors.
- Foreign investors may be attracted by the facility of using LCA services (where they carry cargo) to source spare parts of machinery etc.

The key point about increased foreign investment is of course the long-established argument that foreign investment has higher productivity because of knowledge and other advantages (this idea was probably first expounded in Hymer, S. H. (1960) The International Operations of National Firms: A Study of Direct Foreign Investment, Cambridge, MA.: MIT Press (thesis 1960; published 1976). It is also linked to the argument that the differentials between the productivity levels of entrants and exits from an economy are a key component of changes in aggregate productivity levels.
Trade and markets

Air freight is unlikely to be significant in relation to large-scale sourcing of supplies (although this is subject to empirical verification). However, it can reduce the time, and associated costs, of delivering relatively low volumes of high value components, or key replacement components of fixed capital (such as parts for machinery). In other words, it can facilitate ‘just-in-time’ inventory strategies. This, potentially, may increase productivity.

However, in many – and perhaps most – local/regional economies, the main trade impacts are associated with increasing tourism flows (both inwards and outwards), resulting from reduced travel costs and increased accessibility. The overall net effect depends on a) the balance between inflows and outflows of tourists and b) additionality effects, that is the extent to which these are new flows or merely changes in the transport used previously by tourists making the same journeys. The outcome is a net increase - or decrease – in market size, and this may impact on productivity via scale economies, or stimulating new (more productive) entrants to the sector. A more refined version of this argument would also take into account not only tourism volumes but also the composition of tourism flows and the consequences for market segmentation, and whether this results in net increases or decreases in spending. This may lead to higher sales per employee in existing tourism firms, or to the entry of new (more productive) firms into expanding market segments.

Conclusions

The de- or re-regulation of air travel has had a number of direct impacts, the most spectacular of which has been the growth of LCAs. This has effectively redrawn the map of accessibility and travel costs across Europe, but especially in the UK. This paper has argued that these changes potentially have had unintended impacts on productivity levels in regional and local economies, consequent on changes in flows of labour migration, trade (including tourism), investment and knowledge.

However, LCA activity is highly uneven regionally. LCAs tend to emphasize the following criteria for prospective route selection: long term cost minimization; availability of efficient facilities; and geographic, demographic and strategic considerations. The number of businesses in a regional economy in the early growth phase is another factor in selection (Ryanair (2005): Ryanair Investor Day 4 October 2005, presentation by Ryanair, www.ryanair.com). Those regions deemed ‘not prospective’ by LCAs may lag in the changing map of accessibility, and this may contribute to regional differences in productivity. Interestingly, Graham writing in the late 1990s thought that there had been little change in the regional distribution of accessibility and connectivity. But, subsequently, there has been strong growth in connections between previously unconnected airports, including those in peripheral regions. More recent commentaries have noted significant changes in levels of accessibility, but at the same time a tendency for greater differentials. For example, Bowen (2002 :425) writes that ‘...the deregulation of the airline industry has tended to reinforce the disparity in access among gateways in global airline networks.’ It is also true that there are problems of non-contestability (monopolies) on some routes. Papatheodorou (2002: 387, emphasis added) considers this to be a significant issue, and argues for new forms of re-regulation to counter non-contestability:

‘The competition authorities should face the failure of contestability conditions in peripheral destinations through price re-regulation of the monopolies where services are sustainable and if such a decision is approved by the European Commission. To
avoid the side effects of the previous system, price regulation should explicitly provide productivity incentives. UK utilities operate relatively successfully under this regime .... and these principles (e.g., price caps, efficiency adjustments) can be applied to aviation .... If such re-regulation is precluded, the national authorities may alternatively introduce tendering providing a monopoly air transport operator’s license to the lowest bidder. Where services are unsustainable overall, the winner would be the carrier that accepts the lowest subsidy’.

While recognizing the importance of non-contestability on some routes, the overall effects of de-regulation have been to increase absolute levels of accessibility in many, and probably most regions, at least in the UK, even if relative differentials may have widened. There is limited evidence available from LCAs, and from consultancy reports, as to the impact of low cost air travel on passenger volumes and – exceptionally – on different market segments. LCAs are changing travel habits, and leading to additional new growth. The European Low Fare Airlines Association argues that only some 37% of its passengers originated from shifts within the airline market, while 59% are new demand. Of the latter group, 71% would not have traveled at all and the remainder would have used alternative means of transport Skeels, J. (2005). A KPMG survey on Hungary, demonstrated that, as a result of LCAs, new traveler segments have emerged in Hungary: about one quarter of leisure travelers and 8% of business travelers would not have traveled otherwise KPMG (2005): Low cost airlines – What's in it for tourism?, KMPG: Budapest. Of course LCAs potentially increase inward and outward flows of tourism, and most of the available surveys provide less information on these. The fragmented, but sometimes insightful, empirical evidence on these issues, will be included in a subsequent report on this component of the regulation and productivity project.

There have also been some attempts to quantify the impacts of LCAs or airport expansion on regional and local economies. These have usually been based on multiplier or input-output analyses. For example, a recent study prepared for the Chamber of Commerce and Industry of Carcassonne-Limoux-Castelnaudary in December 2003 showed that the 253,000 low fares passengers who visited the region during one year generated €8.4 million of direct income for the region. The indirect monetary flow, including all tourist expenses and real estate investment, reached €135 million. The so-called “induced flow”, i.e. increased spending generated by those employed in industries affected by direct and indirect impacts, amounted to €272.4 million (Ract Madoux – Groupe Second Axe, Rapport d’etude: Impact socio-economique de la compagnie aerienne Ryanair dans la region et alentours de Carcassonne, December 2003. There are similar estimates for other airports and for some airlines.

While there have been estimates of the aggregate effects on employment, spending or investment, there are not – to the best of our knowledge – been any attempts to analyse the productivity consequences for local economies. A small-scale scooping study lacks the resources to undertake such an empirical analysis. However, this could be approached in a further research in various ways:

- An aggregate level study of regional economies within the UK (or the EU) which relates changes over time in regional productivity levels to air travel (including LCA) services. This could also seek to examine the different levels of regulation, although it may be difficult to quantify these.

- More detailed passenger surveys which established the additionality affects of new air services (including LCAs) in relation to different air travel market segments. A comparative study would allow potential differential regional economic impacts to be identified.
• Firm-level surveys to identify their use of regional air services (including LCAs), and evaluation of the contribution of these to firm performance. This could be related to productivity measures for individual firms.

• Qualitative case study research to explore how practices (employment, innovation, out-sourcing etc) within individual firms are mediated by the existence of increased air services (including LCAs) as a result of de- or re-regulation.

Ideally such research would be longitudinal, focusing on a period of significant change in LCA services within a particular region. In practice, there are considerable challenges in isolating unintended impacts of de- or re-regulation of air travel, and in operationalising a research design to estimate these. However, further research at both the conceptual and the empirical level would help to identify the range and nature of the unintended consequences of de- or re-regulation of air travel. Any such research would also need to focus on the spatiality and temporality of these consequences, i.e. the way that national and EU regulation are played out differently at the regional/local level, and over different time horizons.

References

Massey, D. (1994) Place, Space and Gender, Minneapolis: University of Minnesota Press
UK CAA (2004): The Effect of Liberalisation on Avi
UK Civil Aviation Authority, Regional Aviation. The Impact of Low Cost Airlines, Presentation given at the 10th World Route Development Forum, October 2004
Moving to the Next Level: A Scoping Study

Beverley Randle, Delores Annon Higon, Elizabeth Watson, Guiliana Battisti, Jillian MacBryde, Marek Szweczyzewski, Neil Burns

Background and Focus of the Study

Government reports (Porter and Ketels 2003; Cox 2005) and the press (Smith 2006; Giles 2006) have highlighted a growing concern about the UK’s productivity performance and competitiveness. “The Competitiveness Report” published by Porter & Ketels (2003) highlights the need for the UK to transition to a higher value economy. This report has triggered further research into what can be done to improve the UK’s situation. One of the reports produced after Porter and Ketels was “Post Porter: Where Does the UK Go From Here?” (Birdi et al. 2003). It states “In order to remain competitive and consistently increase profitability, firms need to move up the value chain in their industries over time” (p21).

The growth of globalisation means the UK has competition that “is as intelligent, knowledgeable, innovative as ourselves except more committed and hungry” (Bititci 2005) it has also altered the value chains of companies as, with most changes, it brings positive and negative outcomes for UK manufacturing companies. It means increased competition for the UK but also a chance for growing economies to develop their manufacturing and service industries (Kaplinsky and Morris 2001). Ultimately as Bititci (2005) points out we are to blame as we as customers demand value for money.

Porter and Ketels (2003) identify the UK’s productivity gap in relation to the USA, France and Germany but there are countries which are moving up the value chain, such as India and China (Lyons 2006). The Cox Review of Creativity in Business (2005) highlights that India and China are not just challenging at the low value-added end but at the higher end of technology and skills.

Finbarr Livesey’s study, Defining High Value Manufacturing (2006), gives some hope in that there are pockets of UK manufacturing delivering high value but the challenge is to make it more widespread.

This paper considers what UK manufacturing firms can do to help reduce the productivity gap at a national level by moving up their value chain, increasing their productivity by generating more value.

“How Can Firms in the UK be Encouraged to Create More Value?” (Edwards et al. 2005) puts forward several strategies for firms to increase their value add. The three strategies are:

- Adopting promising practices
- Creating value - Product, process and service development. New products or services that customers are prepared to pay more for or the new offering attracts more customers.
- Re-positioning along the value chain

The focus of this literature review is the third strategy, that of re-positioning along the value chain. The researchers believe that there is less understanding of this strategy and further
work is required on this subject. The other two strategies have been well covered in the literature on strategy, marketing, innovation and general management literature (for example Nick Bloom, John Bessant etc).

At the outset the literature review aims to map the current literature in the area and answer some of the initial research questions (including how do firms reposition in the value chain? and is it worth it?) and identify gaps in knowledge for further investigation.

**Literature Review Methodology**

The research team adopted a systematic approach to reviewing the literature (Leseure et al. 2004). The process of the systematic review started with identifying the need for the review, selecting a review protocol, identification of key words (value chain, manufacturing, supply chain and innovation) and defining the scope.

The focus of the review was agreed to be an examination of the literature that discussed the strategy of repositioning along the value chain. The initial research questions for the study were:

- What is the value chain?
- What are the ways for companies to move in the value chain?
- Do companies that move up the value chain get the productivity gains they expected?
- What are the links between productivity performance (at the firm and national level) and applying the strategy of moving up the value chain?

Figure 1 illustrates the discussions of the research team had on the scope of the literature review.

<table>
<thead>
<tr>
<th>It is</th>
<th>Is not</th>
</tr>
</thead>
<tbody>
<tr>
<td>To establish what has been done already</td>
<td>An evaluation of value chain models</td>
</tr>
<tr>
<td></td>
<td>An in depth analysis of the literature on productivity</td>
</tr>
</tbody>
</table>

Figure 1 – Scope of the Literature Review

Figure 2 shows the process followed to build up the list of papers to review. The keywords (value chain, manufacturing, supply chain and innovation) were identified by the seven members of the research team and validated by the principal investigator.

The research team followed the methodology for ranking the articles used by Leseure et al. (2004). From a list of 62 articles the research team were asked to grade the articles’ titles and abstracts by the following criteria of A (should be in shortlist), B (uncertain), or C (should not be in shortlist). The approach deviated from the Leseure (2004) approach as the rating A, B or C from the four members of the team whom scored papers were converted to A=3, B=2, C=1 and a total for each reference given. Papers scoring 9 or more were considered for the literature review. There was additional references used that were identified by the principal investigator after the scoring of the original references.
The Productivity Gap and Measures of Productivity at a National and Firm Level

The productivity gap

Before proceeding onto the main focus of the literature review, re-positioning in the value chain, the research team had some preliminary discussions on how this strategy ties up with the AIM research theme of productivity. There was also a level of pre-understanding work that was required to ensure the research team had a common understanding. This called for a basic review of the literature on productivity.

At a high level, productivity is the amount of output for input used (ESRC 2004) and is a measure of competitiveness (Porter and Ketels 2003). There has been a lot of publicity and research generated about the UK’s productivity gap with that of the USA, France and Germany but the exact nature, size and cause of the UK’s productivity gap has generated different opinions.

Publicity has included a recent BBC Radio 4 In Business programme entitled “Not Very Productive” (4 May 2006 20.30hrs). Peter Day reported that when the UK’s productivity was considered in terms of output per head it was found to be similar to that of Europe but the UK’s employment rate (labour force utilisation) was higher than in Europe. Furthermore the UK’s output per head is less than that of USA.

The ESRC’s Seminar Series on the subject of the UK’s Productivity Gap (2004) offers some explanations; it states that the gap between the UK, France and Germany has been caused by a lack of skills and investment. The USA is still ahead after discounting the effect of skills and investment; this remaining difference is put down to differences in ways of working.

To understand what can be done at the firm level to improve the national productivity performance the two levels will first be considered individually.
**National Level**

A measure of economic activity at the national level is Gross Domestic Product (GDP). According to Porter & Ketels (2003) the UK’s GDP performance per person is influenced by the following factors: labour force utilisation - the amount of people and time contribution to the labour force and labour productivity - output per hour worked. Porter & Ketels (2003) point out that the opportunity to reduce the UK’s productivity gap lies in labour productivity, as labour utilisation is high and therefore not an issue. Although not in the national interest, if unemployment was higher (labour utilisation was less) it is possible that productivity would increase (BBC Radio 4, 4 May 2006, 20.30hrs).

**Firm Level**

The level of the firm is the unit of analysis for this study. We therefore needed to be clear on how productivity could be measured at this level. Labour productivity is the most commonly cited measure of productivity (ESRC 2004) and has been identified as one of the key performance indicators for the National Manufacturing Strategy.

The research team agreed that Value Added (VA) per employee (Beachman 2002c) could be used as an indicator of labour productivity at the firm level. Using value added has some limitations, as it cannot be used in isolation as indication of a firm’s overall performance. It also provides only a snapshot of the firm’s performance.

Cox (2005) offers the idea that creativity can be a key to greater creativity. The London Development Agency (2005c) produced a report into the production industries in London that stated

“To improve productivity, a manufacturing firm must increase the value of the goods it produces per employee….Success depends not just on cutting costs, but on continually creating goods and services that people value and want to buy.”

It is not the purpose of this paper to enter into a debate on the measures of productivity but rather to study the literature on a strategy that can potentially improve the productivity performance of a firm.

**Value and its Many Meanings**

In the early stages of the research, there was also a need to clarify our understanding of value given the multi-disciplined background of the research team. The subject of value offered a good foundation before moving onto the more complex idea of the value chain. Again we turned to the literature to clarify the meanings.

Value has various meanings and there is agreement in the literature (Livesey 2003b and Ramsay 2005) that the definition varies depending on the discipline using it and to whom the value is being considered in relation.

The fields considering the definition of value include economics, marketing, strategy and operations (Ramsay 2005). Value can be different for employees, investors, owners and the country as a whole (Livesey 2006). Of these stakeholders the customer is the dominant focus in the majority of the definitions of value in the literature.
The traditional “engineering” view of value (value analysis, value engineering etc.) tends to look at ways of maximising the functionality whilst eliminating waste. This view is still seen today in engineering disciplines, with authors such as (Womack and Jones 1994) encouraging companies to focus on the whole rather than the parts, thus allowing companies “to differentiate value from waste”.

To Porter a firm is profitable if the value it commands exceeds the costs involved in creating the product (Porter 1985). (Merrifield 1991) defines value as the increase in value that occurs at each stage of the manufacturing process and value resides in the concentration of resources focussed on selected business areas. Moving from production towards exchange, (Condra 1985) interprets value as a fair return in goods, services or money in exchange for some similar (competitors’ product). (Treacy and Wiersema 1996) go further, defining value as resulting from the fulfilment of customers’ expectations through which the firm achieves the economic benefit. (Miles and Snow 1978) say value comes from choosing customers and narrowing the operation focus to best serve that market segment; customer satisfaction and loyalty does not, by itself, create unmatched value. Livesey (2003a) has more recently offered an activity based view of value. This states that companies provide value through products and services based on the choices they make in the arrangement and selection of the activities they carry out.

It is from the strategic management literature that the seminal contribution on the strategic value creation process has been developed, based largely on the works of Michael Porter and the concept of the value chain (1985). This permitted marketers to think beyond categories of perceived value to the strategic means and processes for delivering to or enabling the customer. This has more recently led to a focus on value in the context of the relationships that exist between suppliers and customers.

Another concept used in the literature is that of value propositions. (Treacy and Wiersema 1996) suggested that the firms compete on three bases – or “value propositions”, namely product leadership, operational excellence and customer intimacy. (Martinez and Bititci 2006) developed this further. Their “value matrix” builds on earlier value propositions developed by (Treacy and Wiersema 1996). This matrix generates six, as opposed to three value propositions which are identified as: Innovators, Brand Managers, Price Minimisers, Simplifiers, Technological Integrators and Socialisers.

These “value propositions” were initially designed to help companies to understand the basis on which they compete, and therefore the strategic and operational issues surrounding their competitive standing. More recently Bititci (2005) has suggested that the concept of the value proposition could also be used by organizations as a means to consider repositioning within the value chain. He suggests that there are many companies who wish to get out of the “price minimiser” situation – so effectively they need to reposition themselves in the value chain. We will come back to this topic in the section ‘What do we mean by moving up the value chain’?

The expectations of value that the customer is looking for is increasing in terms of time to market and need for innovative and customised products, making value capture harder (Livesey 2003a). That is based on the principal that the manufacturer has the capabilities to exploit the value. In the example of the pharmaceutical industry, new companies with the
technology capability have the ability to exploit the value while the more established
competition have to acquire the relevant capabilities in order to compete (Champion 2001).

The concepts of value that are of most use when considering how to reposition in the value
chain are the value proposition, the value chain and the activity based view. The value chain
offers a means of considering all the activities involved in delivering a product or service to a
customer, without the boundaries of firms or countries. This will be explored in more detail
in the next section.

**Value Chain**

What is the value chain? In the literature there is a lack of consensus on the meaning of the
term “value chain”. The descriptions vary from a tool for analysis (Sturgeon 2001), a
business system (Walters and Lancaster 2000), a description of activities (Kaplinsky and
Morris 2001) or a method of facilitation (Normann and Ramirez 1993).

Other definitions of the value chain include the idea that the value chain is a sequence of
value-added activities (Sturgeon 2001).

Kaplinsky and Morris offer an easy to understand activities based description of the value
chain:

“The value chain describes the full range of activities which are required to bring a product or
service from conception, through the different phases of production (involving a combination
of physical transformation and the input of various producer services), delivery to final
consumers, and final disposal after use” (Kaplinsky and Morris 2001).

Walters and Lancaster (2000) offer a definition based on preceding literature that is harder to
related to than the Kaplinsky and Morris definition: “A value chain is a business system,
which creates end user satisfaction (i.e. value) and realises the objectives of other member
stakeholders” (Walter and Lancaster 2000).

The value chain has also been expressed in the form of models to aid understanding but there
are two factors that affect the value chain and its successful operation that are not visible in
the value chain model. These are information management and relationship management
(Walters and Lancaster 2000). The main types of value chain models are outlined in the next
section.

**Value Chain Models**

Value chain models range from the simple that resemble a supply chain with the suppliers on
the left and the customer on the right to models with the customer as the first step in the chain.
An example of a model with the suppliers to the left and the customer to the right is the Porter
model (1985) as illustrated in Figure 3.

A more recently published model by Walters and Lancaster (2000) puts the customer as the
first actor in the value chain as illustrated in Figure 4, showing the customer first with the
value proposition as the interface between customer and firm.

As the intention of the scoping study is to look at movement in the value chain and not to
evaluate the models of the value chain, the analysis will remain at highlighting the different
types of value chains.
Although the value chain is not easy to define and not without its opponents such as Ramsay (2005) who criticises it for being entirely customer focused and not existing beyond the literature, it is not necessary to carry out an evaluation of the definitions and the different models as it would not directly aid the understanding of how companies can reposition in the value chain. The need to understand how to reposition in the value chain is more important because some activities in value chains add more value than others and therefore are more lucrative (Schmitz 2005). Birdi et al. (2003) illustrate this point with an example of personal computer manufacturing where there are several firms involved in design through to assembly but critically not all of them add an equal amount of value to the end product.

Understanding that there are positions in the value chain that offer the potential for manufacturing firm to generate more value if they are not already a high value manufacturer brings us to the subject of repositioning in the value chain.
Moving in the value chain

Moving up the value chain has been billed as an “essential strategy” for UK firms to ensure that they can be competitive and ultimately survive (Birdi et al. 2003). This strong recommendation for the strategy is not surprising given the age of the heavily vertically integrated manufacturing firm is gone (Champion 2001, Livesey 2003b). Also there is a perception that value is moving in the value chain (Wise and Baumgartner 1999). This all adds to the need to understand how firm can move in the value chain.

There are several terms used in reference to moving in the value chain. These are:

- Moving up the value chain
- Moving down the value chain
- Moving upstream in the value chain
- Repositioning in the value chain
- Moving downstream in the value chain

The most common of these terms, from an internet hit survey carried out at the time of writing, is “moving up the value chain”. This may be due to the positive connotations of the phrase rather than being in reference to a specific direction of movement in a chain.

What do we mean by moving up the value chain?

Moving up the value chain is a term often used but lacks a clear established definition. This section examines the existing descriptions of moving up the value chain and offers a new definition.

One of the problems we found with the term moving up the value chain relates to the different perspectives on value. We can see at first glance that the definitions offered by Edwards et al. (2005) and Bititci (2005) look similar. However there are fundamental differences.

Edwards et al. (2005) offers the following definition, “Fundamentally change position in the value chain – moving to a position where their products and services inherently generate more value.” Bititci (2005) defines moving up the value chain as “to produce products and services that the customers are prepared to pay more for, thus changing the basis of competition from cost to something else.”

Edwards et al. (2005) is referring to value in terms of stakeholder return regardless of the move in the value chain relative to the customer. In this scenario it would imply there is a productivity gain by increasing the value added.

Whereas in the definition offered by Bititci (2005) there is no guarantee of productivity gains as you may be able to charge the customer more but it does not take into account the net outcome having potentially generated extra costs to move and remain in that new position.
We propose the following definition based on the literature: successfully moving up the value chain is changing the sphere of activities in the delivery of value to the end customer in a way that increases the firm’s productivity.

Table 1 shows the literature comparison of the different perspectives of repositioning in the value chain and how the strategies compares against the 3 that Edwards et al. (2005) propose for how UK firms can generate more value.

Table 1 illustrates the findings of the literature search that was focused on the strategy of repositioning in the value chain, results included literature that was more in line with adopting better practices or innovating the products or services. Not all of the results mapped to just one of the strategies outlined by Edwards et al. with some fitting across both innovating products or services and fundamentally repositioning in the value chain. For example “Comprehensive Services - the manufacturer providing more services for the customer” (Wise and Baumgartner 1999) is classified under both moving in the value chain and innovating to produce products or services as it depends on the level of change to the firm’s existing value adding activities as to wither it could be described as a “fundamental” move in the value chain or an innovation to existing products or services.

Distribution control (Wise and Baumgartner 1999) could only be classified as moving the value chain if it represented a significant change in (or additional) activities carried out by the firm otherwise it would be classified as “other”.

The literature search was centred on moving in the value chain but some of the descriptions, listed in Table 1, fitted closer to the other two strategies mentioned by Edwards et al. (2005) and did not necessarily involve moving in the value chain.

Livesey (2003) echoes Wise and Baumgartner’s (1999) grouping of ways to move in the value chain, with the additional point about offering the manufacturing capability as a service rather than competing on product production alone. Therefore being an enabler to other firms’ movement in the value chain by being the recipient of the outsourced activities.

An additional means of value chain upgrading that is an outlier to the 3 strategies put forward by Edwards et al. (2005) is the concept of entering a new value chain as opposed to moving up in the existing value chain (Kaplinsky and Morris 2001).

Under the definition of restructuring the value adding activities as a means of moving in the value chain, Edwards suggests outsourcing in the form of changing the value adding activities carried out by the firm. This has to be considered with caution as it is not a guaranteed means to get productivity gains (Edwards et al. 2005).

Another tool for repositioning in the value chain has been outlined in terms of strategies to pursue by Bititici (2005) based on the value matrix as previously discussed in the earlier section on value. These are more general and do not all directly compare to the literature structure in table 1. From the definition of moving up the value chain and the strategies outlined, the suggestion is that companies want to get out of the “price minimiser” category in order to move up the value chain, yet if we take the Edwards (2005) view a company could be successful within the price minimiser category.
Table 1: Literature Comparison Table

<table>
<thead>
<tr>
<th>Author</th>
<th>Perspective</th>
<th>Increasing efficiency and effectiveness through better practices</th>
<th>Innovating to produce products or services that generate more revenue</th>
<th>Fundamentally changing position in the value chain</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Section 4.1 Moving up the value chain</td>
<td></td>
<td></td>
<td></td>
<td>Outsourcing – New make or buy decisions</td>
<td></td>
</tr>
<tr>
<td>Kaplinsky and Morris (2001)</td>
<td>Value chain upgrading</td>
<td>Process upgrading - improving efficiency within each node in the chain and between nodes in comparison to competitor companies</td>
<td>Product upgrading - introducing new products or improving existing products</td>
<td>Functional upgrading - increasing value added by changing activities done within firm or moving them within the value chain</td>
<td>Chain upgrading - entering a new value chain. Example given as moving from manufacturing radios to phones</td>
</tr>
<tr>
<td>Wise and Baumgartner (1999)</td>
<td>Moving down the value chain (nearer the customer)</td>
<td>Embedded Service - the manufacturer providing more for the customer in the product (e.g. Honeywell)</td>
<td>Distribution Control - gaining control of the distribution activities (e.g. Coca-Cola)</td>
<td>Comprehensive Services - the manufacturer providing more services for the customer (e.g. General Electric)</td>
<td></td>
</tr>
<tr>
<td>Livesey (2003)</td>
<td>Manufacturing services to rival business services</td>
<td>Offer services based on goods they manufacture - total care solutions</td>
<td>Offering manufacturing capabilities as a service</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
There are a couple of other points highlighted by the literature that affect a firm’s ability to reposition in the value chain that is ability and power. Even if a firm is able to come up with a strategy for repositioning in the value chain that is not the whole picture as there is the execution of the strategy. The execution of strategy is often the part companies find difficult (Birdi et al. 2003). Firms might find some to the classic barriers to major change such as finance or skills. The firm depending on the position and relationships it holds within the value chain may also find that it does not have the power to move.

The ways of repositioning in the value chain taken from the literature have been summarised in Figure 4 based on the literature.

**Example Firms**

The literature gives examples of firms that have already reached the status of high value manufacturers in the UK these are Cadbury Schweppes, Rolls-Royce and GlaxoSmithKline (Livesey 2006). Wise and Baumgartner (1999) offer a more international list of companies that are thriving having moved closer to the customer. The examples given are Honeywell, General Electric, Nokia and Coca-Cola.

IKEA is a firm used as an example by Normann and Ramirez (1993). It is a good example of firm that allows the customer to carry out some of the activities that a traditional furniture retailer would do, i.e. the assembly and delivery of the furniture. In return IKEA passes some
of these cost savings onto the customer who in return is willing to carry out the additional activities.

Acer computers of Taiwan are a company that has gradually moved toward the customer. Acer started off as a component manufacturer progressing to a brand name in personal computers (Birdi et al. 2003).

The examples given in the literature highlight a gap, most of the firms discussed are recognisable which indicates that there has not been as much written about SMEs.

**Discussion and Gaps**

The research areas of interest emerging from the literature are around the how to reposition in the value chain, why reposition and what changes in order to reposition. The how is a big subject, the why and the what seem simpler areas to start with. Figure 5 shows some of the Drivers for repositioning developed from the literature.

Understanding these questions in particular for SMEs would be worthwhile, as it would lay the foundation to develop strategic tools for helping some of the most vulnerable and significant firms in the UK economy. There are some means of assessing if it is worth moving in the value chain offered by (Martinez and Bititci 2006, Wise and Baumgartner 1999). These tools have yet to be applied and validated.

The published work on the value chain has mainly been written from an operations perspective and there is a lack of applied tools and method studies that include financial analysis to validate the benefit to firms of moving in the value chain. The tools and techniques to help firms move in the value chain exist within the consulting sector but they are not widely published and therefore the empirical evidence is not available to substantiate if it is worth moving in the value chain.

![Figure 5 – Drivers for repositioning in the value chain](image)

The gaps are:

- Strategic tools – there is a lack of proven tools to help firms assess how to move in the value chain and if it is worth moving.
SMEs – given the significance of SMEs (they account for 50% of UK GDP, not just manufacturing (Cox 2005)) and their vulnerability in the value chain (Edwards et al. 2005) there are a few case studies looking at them but little in terms of evaluating how they can be helped to move up the value chain.

There is a gap in the understanding of the connections between what can be done at the firm level to affect productivity at the national level.

**Conclusions**

The literature survey set out to answer the initial research questions and to explore what had been done in the field of moving in the value chain. Some of the initial research questions remain unanswered and require further investigation before they can be answered. The process of studying the literature has also highlighted additional gaps.

Questions answered:
- What is the value chain?
- What are the ways for companies to move in the value chain?
- How is productivity considered at the firm and national level performance?

Questions unanswered and gaps requiring further investigation:
- Do companies that move up the value chain get the productivity gains they expected?
- What are the links between productivity performance (at the firm and national level) and applying the strategy of moving up the value chain?
- Strategic tools – there is a lack of proven tools to help firms assess how to move in the value chain and if it is worth moving.
- SMEs – given the significance of SMEs (they account for 50% of UK GDP, not just manufacturing (Cox 2005)) and their vulnerability in the value chain (Edwards et al. 2005) there are a few case studies looking at them but little in terms of evaluating how they can be helped to move up the value chain.
- There is a gap in the understanding of the connections between what can be done at the firm level to affect productivity at the national level.

The future research agenda should also build on the drivers for moving in the value chain (Figure 5) and the strategies for moving (Figure 4). The finding from the literature can be tested in case study companies and revised depending on the results.

**References**


Intra- and Inter-Firm Knowledge Transfers and Productivity in the Retailing Sector

Dolores Añon Higon, Jeremy Clegg, Irena Grugulis, Allan Williams, Nicholas Vasilakos, Ödül Bozkurt

Introduction

Retailing is a crucial sector and major contributor to the UK economy in term of both output and employment. Recent figures from the British Retail Consortium (BRC) suggest that the UK retail sector generates almost 6% of total GDP. In 2001 retailing generated £49 billion (5.6% of GVA) and employed over 3 million people, about 11% of the total workforce. In 2002 retail sales amounted to £239 billion, about 35% of total consumer spending in the UK. There were 188,870 VAT registered retail based enterprises in the UK in 2003, and the sector was, after housing and education, the 3rd largest services based industry in 2004 (DTI 2004).

Not surprisingly, given the importance of retailing, the productivity analysis of the UK retail sector has become an area of policy interest in recent years. Several comparative studies at an aggregate level have shown a labour productivity gap of UK retailing when compared with other countries, notably France and the US (McKinsey 1998, O’Mahony and de Boer 2002, Van Ark et al 2002). Accordingly, even after allowing for variation of hours worked, productivity gaps remained, comparable scores ranging from 10 to 63 per cent greater in the US and 25 to 59 per cent greater in France (DTI 2004:33, ESRC 2004). These studies have also assessed the extent to which the size of the productivity gap has evolved in the last decade. Despite the different methodologies and datasets employed in each case, most of them seem to agree that whilst the labour productivity gap between the UK and US retailing sector narrowed in the early part of the 1990s, it widened significantly in the latter part of the decade. In addition, Basu et al. (2003) find that retail trade, together with hotels and catering, account for about three quarters of the US productivity growth boost in the late 1990s; and for one third of the UK productivity slow-down during the same period.

The exact extent of the UK productivity gap remains uncertain, however, due to methodological complications in the comparison of productivity between national economies. For example, the Templeton Report, which reviews the three principal recent studies on productivity, recognizes the common findings but also challenges the comparisons on methodological grounds. Accordingly, the three studies, respectively carried out by McKinsey (1998), NIESR, and Groningen, have found significant gaps in labour productivity for the UK retail sector as a whole, compared to the US and France, but the report urges further studies to ensure comparability of figures. There are additional methodological caveats in the study of productivity because the concept involves both outcomes and processes. Research to date has largely privileged highly-aggregated, almost exclusively quantitative approaches that aim to produce quantifiable “scores” that pertain directly to said outcomes. This mode of inquiry is integral to productivity analysis, but stands to be complemented in significant ways with qualitative research that focuses on processes. A methodologically diverse approach is therefore the most appropriate for the analysis of productivity in general, and the claimed UK productivity gap in particular.
While methodological considerations therefore remain highly relevant in productivity analysis, these are directly linked to the effort to identify correctly the specific factors that result in higher or lower levels of productivity. Research to date has suggested a number of such factors, based especially on the different trajectories of UK and US retail productivity over the past decade. For example, Griffith (2003) distinguishes poor management, lower skilled labour force, the inadequately competitive UK retail market, land regulation/planning, and the adoption and use of information and communication technologies (ICT) as the factors behind lower productivity in the UK. Others have offered alternative factors as having primary significance.

In our review, we provide a detailed survey of the growing literature on UK productivity with emphasis on both methodological and substantive issues, and offer a multi-tiered framework informed by this survey for the further investigation of the topic. Importantly, we do this at a specific moment when the sector's development reflects the general global trends of increased internationalisation. The UK economy has witnessed the growing presence of foreign-owned retailers, especially during the past decade. As is the case in other developed host economies, such growth has been driven both by acquisitions, and to a smaller extent greenfield investments (Burt and Sparks 2003:501). The acquisition of one of UK's largest supermarket chains, Asda, by the world's leading retailer, Wal-Mart, at almost $11 billion, is one of the biggest cross-border acquisitions in retail in the world to date (Dicken 2003:501). Although retailers based in the UK still claim dominant positions in their respective segments, often by wide margins, an increasing proportion of UK retail sales are nevertheless captured by non-UK businesses operating in the country (Burt and Sparks 2003:26). The success of retail chains at the “no-frills” end of discount retailing is one specific example that illustrates the trend, however nascent. At the same time, a number of the larger UK-based retailers have embarked on the expansion of their overseas activities. While the proportion of revenues generated from overseas currently remains relatively low for the retail sector at large, it has registered rapid growth for some of the leading practitioners of multinational activity such as the supermarket chain Tesco. It is also the case that the number of retailers that are either considering or undertaking overseas activity is on the rise. Given this transformation of UK retailing, the dimension of multi-nationality must also be incorporated into the analysis of productivity in the sector and in its firms. In fact, the co-presence of multinational and domestic firms in the retailing sector allows for the adoption of useful methodological tool in investigating the productivity gap question through firm comparisons.

In the remainder of this piece, we follow up on these fundamental insights and do the following: We begin by discussing the measurement issues pertaining to productivity in retailing as they manifest themselves in econometric analysis. Secondly, we offer a list of determinants of retail productivity to improve upon such econometric analysis of productivity and explain why each deserves particular attention. Thirdly, we turn to the significance of “knowledge” in the retailing business, identifying two central avenues of investigation that open up the hugely popular but still very ambiguous concept. On the one hand, we consider the various modes of knowledge transfer, including both those inside and between firms, and the significance these hold for the productivity of individual retailers as well as the retailing sector as a whole. This concern with knowledge as a resource to be transferred, thereby informing productivity levels, is especially important against the backdrop of the aforementioned context of internationalization. On the other hand, we question what the firm strategies adopted in order to enhance productivity mean for the knowledge content of the different forms of employment in the retailing sector. The focus on the knowledge content of retail sector jobs along an axis of highly codified to highly personalised is warranted for at
least two major reasons. First, that the UK productivity gap is largely or at least significantly due to a skills shortage in the UK workforce is a popular claim that has helped bring about a series of government initiatives to improve the supply of skills and qualifications (Lloyd and Payne 2006:151). Secondly, a closer look at the real-life processes whereby skills are “applied” and measured in actual workplaces provides a necessary corollary to the econometric analysis of the relationship between the quality of the sectoral workforce and productivity. Finally, we conclude by suggesting directions for future research on productivity in the UK retailing sector.

Measurement Issues in the Assessment of Productivity in Retailing

The aim of this section is to review different concerns in the literature related to problems in the measurement of retail productivity. There has been a long-standing discussion concerning the methodological complications involved in the measurement of productivity in the service sector in general, and of distribution and hence retailing in particular (Cox 1948, Goodman 1985, Borin and Farris 1990, Reardon and Vida 1998, McGoldrick 2002, among others). The Templeton Report (2004) also notes two fundamental reservations with the results offered by the aforementioned comparative studies. The report argues, first, that other inputs, and not just labour, contribute to the generation of retail value added. These inputs include IT systems, capital investment in shops and the space they occupy, and the different combinations of capital, labour and land used by retailers in the countries included in the studies. Secondly, the Templeton Report raises questions about the accuracy (mainly due to data limitations) of the estimates used in calculating labour inputs across countries. This input varies in quality and kind, due for instance to variation in skill levels, which in turn are defined to a significant degree by the variation in the quality of part-time and casual staff. Variations in hours worked likewise complicate a direct comparison of labour inputs across different countries. These issues will be discussed to a greater detail in this survey.

Productivity Measurement in Retailing

A definition of retail productivity, provided by Jefferys et al (1954): “An increase in productivity in distribution can be described simply as:

Either the provision of the same output, that is the same group of goods, and the same volume of services to the consumer with a smaller input, that is at a lower unit cost as measured by the outlay of the factors of production, employment, space, capital etc.
or the provision of increased output, that is the same amount of goods a greater volume of services to the consumer with the same input, that is the same unit of cost as measured by the outlay of factors of production.”

The most commonly used measure of retail productivity is that of labour productivity (see Table 1), that is the ratio between a measure of output (frequently sales or gross value added) and a measure of labour (the number of employees or man-hours worked). Though this measure of productivity is the one that is most often used in retail productivity studies (perhaps due to data availability and ease of construction), it is vulnerable to the implicit assumption of only one factor of production, that is labour. As such, in cases where the importance of factors of production other than labour is not trivial, such as physical capital or land, the use of labour productivity may lead to biased results. Recent empirical evidence
appears to lend support on the validity of this concern. McKinsey (1998) find that while the UK had lower labour productivity in retail in 1995 in comparison to the US and France, its capital productivity was in fact significantly higher in that year.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Productivity Measure</th>
<th>Output Measure</th>
<th>Inputs</th>
<th>Retail Level of Analysis</th>
<th>Sample Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>McKinsey (1998)</td>
<td>L, TFP</td>
<td>Gross Margin</td>
<td>L (hours worked), space</td>
<td>Food retailing</td>
<td>UK, France 1995</td>
</tr>
</tbody>
</table>

A more refined measure of productivity is Total Factor Productivity (TFP, henceforth). Conceptually, TFP should be measured as the ratio between real output and a weighted sum of real factor inputs. The weights should, in principle, reflect the relative importance of each input contribution to production. In practice, one can discern two theoretically distinct methods for computing the index of inputs. These can be distinguished, among others, by the assumptions for determining the weights assigned to the different types of input. The first method, the growth accounting approach, predicts that under some simplifying assumptions, factor income shares should be used as weights, this income shares approximate production elasticities or the effects of a 1% change in individual inputs on outputs. The second approach, the econometric method, weights the different types of inputs on the basis of their relative ability to explain output through regression analysis. Essentially, both methods are constrained by the tangibility of inputs, in the sense that some inputs (such as organizational capital and knowledge) are unobservable and, therefore, by nature difficult to quantify. In theory, the impact of such variables can be accounted for through the use of appropriate (observable) proxy variables, highly correlated with the (unobservable) production inputs (e.g. observable human capital stock within a firm can be used to approximate the stock of knowledge within this firm). In practice, however, the use of appropriate proxies itself is subject to data availability, which, in several instances, is severely constrained, especially at higher levels of data disaggregation.

At this point it is important to notice that the risk of incurring measurement errors is frequently exaggerated by the seasonality of demand that tends to characterize the UK retail sector. The issue of seasonality, when combined with the data collection techniques that are usually employed in the construction of datasets, may give raise to “snapshot” data, not always representative of the actual dynamics of the sector (Reynolds et al 2004). A well documented example is the wide use of casual or part-time employment in retailing:
Apparently, in the absence of appropriate data, the inclusion of casual labour as regular workforce can lead to misleading results: If the data in labour has been collected in a peak season time, the representation of labour in the production process will be over-represented. Similarly, if the data has been collected in an out-of-season point in time, the contribution of labour will be under-represented. Currently most datasets have started distinguishing between casual, part time and regular employment, but seasonality still remains a strong issue of concern as the once-a-year method of data collection fails to smooth out enough the series that are being made available.

Reynolds et al. (2004) raise concerns about measurement errors in commonly reported measures of productivity. In a more general context, these errors may be triggered by issues related with:

Imperfectly competitive retail markets: prices may not reflect quality (or cost) of the service, if the market in which the firm operates is imperfectly competitive – or, in cross country regressions, it may just reflect differences in the degree of competition of each market.

Measurement of inputs: market rigidities (such as excessive planning regulation and strong labour unions – though the second may not be relevant for the UK) may distort the compensation of factors of production from their equilibrium values.

Other issues of more technical nature may include methodological problems related to international comparison and the “aggregation bias”. Being one of the most commonly reported issues of concern, aggregation bias essentially refers to the loss of variation that is imposed on a dataset when the analyst opts (or is forced - due to data availability constraints) to aggregate across firms or sectors. When this loss of variance is not uniform across all independent variables, aggregation may lead to biased estimators, the extent of the bias being determined by the noise that has been imposed due to aggregation. A direct implication of such errors would be the creation of biased estimates of productivity, for which the bias direction would be in line with the direction of the distortion of prices.

A related problem in measuring retail sector productivity may involve the wide heterogeneity of the retail sector. Reynolds et al. (2004) raise the issue of comparing productivity across retail propositions: as long as the retail sector offers differentiated products (e.g. superstores vs. convenience stores), different suppliers may utilize factors of production to different intensities. In their example, they mention characteristically that superstores may be more capital-intensive than convenience stores. Thus if the measurement of productivity is based on labour productivity, the convenience stores may show up as much less productive. This argument suggests that a measure of productivity that averages across factors may be more appropriate to use on that account.

Finally, one of the key measurement issues associated particularly to productivity analysis in service sector is the definition and measurement of real output and inputs. Indeed, the precise definition of outputs and inputs in the retail process lies at the heart of the productivity analysis. Most services, including retailing, have been characterized by non-material outputs, which are partially a reason why the service sector came to be discussed as “low-tech, low-productivity industries with little impact on a country’s economic performance” (Preissl 2000:125). The peculiar idiosyncrasy of retailing, in terms of the intangible nature and heterogeneity of its output, participation of customers in the production process and high degree of simultaneity between production and consumption (Goldman 2001), pose
significant challenges of both definition and measurement. In the following we explore the output measurement issue further.

**Measures of Retail Output**

Early contributions to the debate around the measurement problems in relation to retail output include Cox (1948) and Hall and Knapp (1955), particularly with regards to the difficulties of constructing an index of aggregate output of retailing services. More recent contributions focus on the measurement problems related to the quality and volume of retail output, as evidenced by Reynolds et al (2004) and the references therein.

Reynolds et al. (2004) provide two examples through which mis-measurement of output may result to biased productivity estimates. The first problem relates to the problem of disentangling the products sold by a retailer and the retail services delivered. A typical example is the opening up of more tills in a convenience store to reduce the queuing time of customers. Though this strategy increases customer welfare it will most likely show up as productivity reducing when the earlier definition of (labour) productivity is used, since the opening up of a new till will result to fewer sales per cashier (unless the market share of the retailer increases as a result of that to such an extent as to prevent this). This undoubtedly is related to the problem of quality adjustment. In retailing, where the process of being served is as much a part of the purchase as any product exchanged, quality is harder to assess (Hean et al. 2003). A successful service encounter requires co-operation (and often participation in work) by customers and it is not clear whether increasing the quality of this encounter is an improvement in product specification or simply means delivering to specification (Keep and Mayhew 1999). In addition, most services innovation is linked to “changes in processes, organisational arrangements and markets” (OECD 2000).

A second issue relates to the problems associated to the deflator index. Recent evidence suggests that the common use of CPI (Consumer Price Index) as a proxy for inflation in the retail sector may result to the overstating of inflation. Characteristically, Nakamora (1999) finds CPI to over-estimate US inflation from 1978 to 1996 by 1.4% a year. Overestimation of inflation will apparently result to underestimation of real output and, as such, potentially, an underestimation of labour productivity.

As Timmer and Inklaar (2005) highlight there is no consensus on how to measure output in retail trade for the purpose of productivity measurement. Many productivity studies, particularly at the store level, use real sales (see Doms et al. 2004) or gross margin (Mackinsey 1998), which is defined as sales less the cost of goods sold. At a more aggregate level, and particularly when international or sectoral comparisons are involved, gross output and value added (see for instance O’Mahony and de Boer 2002 and Basu et al. 2003) are more frequently used. The main difference between the two being the purchases figure, which is deducted in the calculation of gross value added.

**Measures of Inputs**

Problems of measuring output are added to by difficulties of defining and measuring the inputs that generate that retail output (Currah and Wrigley 2004). This is true whether the focus is on the labour productivity of the retail trades overall or on total factor productivity (TFP). In the same way that one should be worried about quality adjustments in measures of output, inputs should be equally adjusted. Griffith and Hamgart (2005) rightfully ask, “Is one
worker, or one hour worked, the same as another?” Adjustment for full-time vs. part-time employees and for the skill composition should be taken into account. Indeed, skills have an important impact on productivity, and studies at the aggregate level show that they are one of the determinants of the UK productivity gap. The UK is behind France and Germany in terms of intermediate skills and behind the United States in graduate skills.

Most assessments of UK retail productivity focus on labour, not on TFP, and therefore take little account of the different forms of UK retailing and the different combinations of other inputs such as land, ICT and physical capital. Differences in labour productivity can be explained in part by differences in the intensity of use of such inputs. When controlling for these differences, more accurate estimates of retail productivity may be obtained, whereas some “dark aspects” of the retail sector may become more obvious: McKinsey’s (1998) study, for instance, shows that when other factors of production than labour are accounted for, the productivity gap for the food retailing sector between the UK and other European countries tends to narrow down (and in certain cases disappear). A further advantage of such an approach is that it enables the identification of inefficient use of non-labour factors of production, ICT being one of them. However, more often than not, data and resource constraints do not permit a careful differentiation and full coverage of all inputs.

The measurement of productivity in services, including retailing, is further complicated because the relationships between inputs and outputs are particularly blurred in service exchanges (Currah and Wrigley 2004). Most services have been characterised by non-material outputs, which is one of the factors which influenced the perception of services as “low-tech, low-productivity industries with little impact on a country’s economic performance” (Preissl 2000:125). In retailing, a service is produced and consumed simultaneously. Therefore, in common with other service sector industries, real time quality control is of utmost importance to productivity; this contrasts with manufacturing, where quality control can be exercised after production.

**Methodologies**

Besides problems related to the measurement of outputs and inputs, there are also specific problems associated with the traditional methodologies to measure productivity. Traditional estimates of retail productivity have been constructed using growth accounting, methodology which became popular with the seminal papers of Solow (1956) and his contemporaries. Under this approach, the definitions that have been used for output and productivity are usually meant and designed to serve the analysis of data from the manufacturing sector. Admittedly, such estimates should be used with certain caution and adjusted to account for the fundamental differences between the two sectors.

In contrast to the relatively homogeneous manufacturing sector, retailing tends to exhibit striking intra- and international diversity. This implied heterogeneity is well documented in the literature (for some interesting figures see Dawson (2004)) and, when combined with other features of the retail sector, it serves to highlight the challenges to valid economic analysis that this sector may hide. A vivid illustration of such a source of intra-national retail diversity can be found within the synthesis of the UK retail population. Given that the retail sector tends to be characterized by much lower barriers to enter and exit than manufacturing. The accuracy of this statement may depend on the extent to which the retail sector is regulated: providing the excessive planning regulation in the UK, barriers to enter or exit may not be as low as in the US, but still in most of the cases tend to be lower than for
manufacturing (mostly due to lower capitalization requirements), it tends to attract a large mass of small-medium sized and family enterprises, managed and staffed by individuals who may pursue very diverse objectives (e.g. self-employment, flexible working hours etc). Measuring, therefore, productivity as volume of sales or profits per worker and assuming a strict profit-maximizing behaviour, may lose many of the other dimensions and entrepreneurial drivers of the retail sector.

Similarly, at the international level, an increasing volume of studies points out to the fact that, opposite to the manufacturing sector, the structure and nature of retailers may differ significantly across borders, as firms tend to adjust their output-service mix to reflect the peculiarities of the local market (Kamakura et al 1996).

An alternative approach to growth accounting is the econometric approach. The standpoint of the econometric analysis to productivity measurement is the estimation of an explicitly specified production (the primal approach) or cost function (dual approach) with the objective of establishing the direct linkage between productivity and the key characteristics or parameters of these functions. The estimated parameters of the underlying production or cost model then are used to derive an index of productivity growth. One important benefit from this approach is that it allows for the careful testing of various features of a postulated model. This is preferable to imposing these features a priori. All these possibilities come at a cost, however. The regression approach is not devoid of problems. For instance, the accurate specification of the functional form and estimation of the parameters of these functions are considered to be crucial to the measurement of TFP growth (Nadiri 1970). Any misspecification or errors in estimating the production or cost function will spill over to the measure of TFP. According to Hulten (2000) there is no reason why the econometric and the index number approach should be viewed as competitors and he quotes examples of synergies that proved particularly productive. These arise in particular when econometric methods are used to further explain the Solow residual.

**Determinants of retail productivity**

Despite concerns about the measurement of productivity, recent empirical studies have show a significant productivity growth in the UK retail sector. Reynolds et al. (2004) find that the UK retail sector account for a significant share of the UK productivity gap with its major international competitors. Similar results are found by Griffith and Harmgart (2005) who found that productivity in the UK retail sector is not just low but has grown slowly over recent years when compared to the US. Recent research has tried to identify some of the factors responsible for the productivity differential (Griffith and Hamgart 2005, Haskel and Khawaja 2003, among others). In the following, we relate some of the factors that have been identified as responsible for the UK productivity disadvantage, with a particular focus on the retail sector.

**Competition and the composition effect**

According to this view, the change of an industry’s productivity can be triggered by two events: The change in productivity of the incumbent firms or by a composition effect through which low productivity firms exit the market and are being replaced by (new) higher productivity competitors. Empirical studies for the US, show that all retail productivity growth in 1990s can be accounted for by the replacement of much unproductive retailers by high productive competitors (see Foster et al. 2002). Similar findings are reported by
Bartelsman et al. (2003), concluding that entry and exit of firms and changes in market shares make an important contribution to aggregate productivity growth.

Although it appears to be a common consensus that the study of entry-exit rates may be central to understanding the evolution of retail productivity measures, severe data limitations, mainly stemming from the level of data disaggregation (shop level) required to perform such exercises, have till recently severely constrained the ability of researchers to evaluate the impact of the competition and composition effect for the UK retail sector. At present, data for the UK is available from the IGD for supermarkets only and has been exploited by Griffith and Harmgart (2005), who identify a switch of investment strategy of the four largest supermarkets from large edge-of-town supermarket centres to convenience (city) stores. At the firm level, Haskel and Khawaja (2003) find that although annual entry rates in the UK are in fact much higher than in the US, the contribution of new stores to aggregate productivity growth is much lower in the UK than the US. They also find larger retailers to have higher labour productivity but, as one would expect, growth in labour productivity is fastest amongst the smallest retailers.

One would expect high entry rates to have a positive impact on productivity growth. The expected positive correlation between the two variables can be justified on the grounds of intensified competition, driving out poorly performing shop and further technological sophistication of new entrants, given that the adoption of new technologies may be easier when opening up a new store. Data for the UK (see Haskel and Khawaja 2003) show that annual entry rates for the period 1998-2000 were between 8.7 % and 10.2% across all retail sectors, while exit rates particularly between 1999 and 2000 were higher. These differences may raise concerns on the degree of market concentration in retailing. Quoting Dawson (2004) provides a good summary of the related statistics. “the proportion of sales accounted by the 10 largest firms in the UK increased from 30% to 38% between 1990 and 2000 and would appear now to be circa 40%. Though generally the same picture seems to be shared in other European countries too, “there are notable differences, particularly in the extent of concentration and the mechanisms of new formation” (Charke et al, Eurostat 2001). In particular, alarmed by the growth of globally organized retailers such as Wal-Mart, Carrefour and Tesco, industrial organization scholars have analyzed the emergence of these large scale retailing, particularly in the form of hypermarkets and the associated increase in market power UK evidence suggests that size does seem to matter in retailing, since largest retailers are found to be more productive than the smaller ones (Reynolds et al 2004) (Dobson and Waterson 1999, European Commission 1999, Competition Commission 2000).

**Planning Regulation**

On the contrary, the extent and intensity of planning regulation should be expected to be negatively related with retail productivity. First, planning regulation might result in retail stores operating below the minimum efficient scale, thus leading to lower productivity levels. A typical example is when regulation affects wages and/or bargaining power of factors of production. Secondly, planning regulation may hinder the opening of new stores or, most frequently, the closure of old ones, thus impeding the substitution of least efficient from more efficient retailers in the way that was described earlier. In fact, recent research by Griffith and Hamgart (2005) suggests that planning regulation did have a significant impact on market equilibrium outcomes and particularly it acted as an entry barrier, although the economic magnitude of the effect may be overestimated.
Whilst land-use planning towards retailing in the 1980s allowed decentralized activity, since the early 1990s there has been a growing consensus on the tightening of restrictions on off-centre and green-field development. Thus it has become much harder to obtain planning permission for developments away from existing town centres and newer forms of retailing such as factory outlet centres and regional shopping centres have become harder to accommodate. Flath (2003) reports evidence for the UK providing a less competitive retail business environment than the US and most other European countries. Following Flath’s, store density is found to be similar to that of the US, but since car-ownership and average travel distances to stores is much smaller in the UK, this could still be in line with a less competitive UK environment. On these grounds, entry may lead to lower prices and better quality.

**The Role of ICT**

Another major determinant of productivity is the technological sophistication of retailers, that is the investment on and use of Information and Communication Technologies (ICT). As appears from a recent study, the trade sector has been transformed over the past three decades from a low technology sector focused on distributing goods to one that makes extensive use of ICT (McGuckin et al. 2005). In theory, ICT investment should be expected to have an inverted-U impact on labour productivity, causing an initial drop in labour productivity at early stages of its implementation and recover later, as the workers become familiar with the new equipment.

The potential importance of ICT on retail productivity growth has recently attracted a considerable volume of intellectual investment, with most of the empirical studies confirming a positive correlation between technological sophistication and retail productivity in US and Europe. More specifically, Doms et al (2004) show that growth in the US retail sector over the 1990s involved the displacement of traditional retailers by sophisticated stores introducing new technologies and processes. Broersma et al. (2003) in a study of the distribution sector in Netherlands show that computer investments had a positive impact on productivity and that the effect was greater in retail than in wholesale trade. Similar conclusions were reached by Hempell (2002) in his study of the effects of ICT investment on business-related and distribution enterprises.

Cross-country productivity studies particularly highlight the large post-1995 productivity gains in the US that are attributed to increased IT usage mainly in the distribution sector. Some even identify a “retail revolution”: For instance, Van Ark et al. (2002) and Basu et al. (2003) argue that the US-UK productivity differential can be partly explained by differences in ICT investment, highlighting how technology started to become a major component of a traditionally labour-intensive industry. More specifically, Basu et al. (2003) argue that slower TFP growth may be seen as a short run adjustment cost, stemming from the diversion of firms’ available resources to reorganisation and learning. They also note that US ICT prices fall faster than UK ones, so that the UK’s ICT and capital investment costs will grow more rapidly for otherwise similar kinds of investments, with a correspondingly differential effect upon productivity. If then retailing becomes an intensive ICT-using sector, one may expect that differential to be considerably higher for this sector. They conclude that this price differential provides US retailers with a competitive advantage over their UK counterparts, contributing to the creation of the productivity differential between the two countries.
The role that the extension of ICT use in retailing has for productivity has immediate bearing on the relationship of the “quality of labour” to the same. We therefore address first, from an economics perspective, this latter relationship. In the final section of our review, however, we highlight some of the potential shortcomings of a purely econometric analysis of the role of labour in total factor productivity in the context of retailing sector, given the complex nature of embodied knowledge in the context of the service sector and of service exchanges.

The composition of employment: the role of skills

The importance of the “quality of labour” has been elaborated extensively from a theoretical (Romer 1990) and empirical perspective (Barro 1992) in economics. We therefore first review the way the composition of employment and the role of skills have been conceptualized in this literature.

Labour specialization through in-house training, education or learning-by-doing seems to be positively correlated with productivity growth. Empirical studies for the UK tend to suggest that the skills gap is a likely contributor to the productivity differential, particularly with respect to France and Germany (Brodberry and O’Mahony 2004). However, most of the empirical studies on the contribution of skills to the UK productivity gap have been focused on an aggregated level and particularly on the manufacturing sector.

Interestingly, despite the fact that the re-opening of the skilled-labour driven - “new growth theory” traces back to the early 1990s, it has been only recently that the debate on the importance of skills has shifted towards the distribution sector. Interesting, indeed, but not surprising, given that traditionally technological sophistication was considered to be more of a privilege of the manufacturing sector and much less of the “deskilled, traditional, labour intensive service sector”. Recent empirical evidence suggests that this picture is eventually changing (Haskel et al 2005). The changing structure of retailing towards a computerized, technological sophisticated sector revives the importance of skilled labour as a driver of productivity growth. At the same time, UK retailing is accused of suffering from poorer quality of labour than their international counterparts (Dawson 2004). Moreover, recent evidence suggests that the composition of employment, that is, the relative proportions of full time vs. part-time or casual workers, may also differ significantly across countries (Reynolds et al. 2004), with the UK employing more part-time workers than US or France. This finding raises further questions about the quality of the labour force in the UK retail sector, and the extent to which current sector-level employment strategy may have a negative effect on productivity growth.

Despite the potential and intuitive implications of the quality of labour for retail productivity, the empirical evidence that has up to date been reported to support this view is rather limited (Haskel et al.2005). This is partly attributed to the lack of available secondary data at firm and plant level, recent evidence suggests that, in the absence of reliable data for skills, the use of wages may constitute the second-best, given the high degree of correlation between skills and wages. Proxying, therefore, level of skills by wages paid to employees may facilitate to overcome this problem enabling the analysis to be carried out at a relatively low level of aggregation (see, for instance, Galindo-Rueda and Haskel, 2005), and to measurement errors and poor quality of existing data for employee qualifications and skills when moving up to higher levels of aggregation, which, when used, tends to produce misleading results in the sense of wrong signs or statistical insignificance. We therefore discuss further below some of the more specific problems pertaining to the exploration of the importance of skills and skill
composition for productivity in retailing and why econometric and qualitative methods need both be adopted in obtaining a robust understanding of the “human element” in retail under separate heading.

**The role of Multinational Enterprises (MNEs)**

Retailing is an appropriate sector for the investigation of productivity in the UK not only because of the overall weight it carries in the national economy but also because it is increasingly internationalised. Although still few studies address international retailing (Palmer 2005:23), recent studies have begun to incorporate the distribution, and particularly the retail, sector in the globalization debate (Wrigley 2000).

The “global turn” in retail studies is precipitated by the fact that, although the retail sector globally retains a domestic orientation in general, recent years have registered a clear and definite turn towards internationalisation. Cross-border retailing witnessed a highly visible surge in the 1980s and 1990s (Godley and Fletcher 2001:31), leading some scholars to note that “the recent wave of internationalization is unique in its scale” (Treadgold 1988:8). The scale of cross-border merger and acquisition activity in the retail sector during the 1990s has surpassed that in all previous times. Both the number of cases and their aggregate value have increased dramatically during the second half of the 1990s (Dicken 2003:501). The rapid growth of trans-national retailing starting in this period is illustrated by the fact that while there were no retail firms among the world's top 100 corporations in 1993, four firms had entered this list by 1999. Of the leading 20 international retailers in 2000, 12 were involved in predominantly food retailing. (Dicken 2003: 496). A relatively small “elite” group of food retailers with extensive—and fast-growing—trans-national operations have played an especially important role in the internationalisation trend (Dicken 2003:497).

Within the UK context, research on the internationalisation of retail is still only nascent (Burt and Sparks 2003, Hallsworth 2002). A much larger body of work exists on the activities and management of multinational corporations either based or operating subsidiaries in the UK, mainly focusing on the diffusion of employment practices, in the context of manufacturing. Various empirical studies have looked at UK-based multinationals and their overseas subsidiaries and inquired whether country of origin effects have been strong in their operations. These studies range from in-depth case studies of firms, such as two in food and automotive components manufacturing (Edwards, Rees and Coller 1999), to the analysis of more encompassing surveys, such as Edwards' study including a range of manufacturing MNCs in chemicals, medical equipment and plastics, engines, beer and spirits, packaging, paint, and aerospace (Edwards 2000). Other studies have considered the hypothesis about multinational corporations' acting as “enclaves” within a different national economy. These studies have included the UK subsidiaries of MNCs based in Germany (Tüselmann, McDonald and Heise 2003), of US-based multinationals in IT, consumer goods, household products, engineering contracting and mechanical engineering in the UK (Ferner et al 2004); and Japanese automotive manufacturers again in the UK. These studies have not yielded a consensus on whether the employment relations at the plants reflected a diffusion of practices from the country of origin or if they more closely matched those at comparable firms in the same locations. The variation in the findings strongly suggest, among other things, the possibility that a “multi-nationality effect” may be largely dependent on specific sectors and the way competitive practices are shaped in them. To date in-depth comparisons of multinational vs. domestic firms in the UK retail sector remain wanting.
As aforementioned, the British retail sector has recently seen a considerable degree of internationalisation, both in terms of foreign-owned firms coming into the British market, and in terms of UK-based firms expanding their operations overseas. Godley and Fletcher (2001) note that, while foreign entrants in British retailing can be traced back as early as 1885, and that there were, on average, over 40 new entrants per decade until 1980, numbers have indeed shown a striking rise after 1990. From 52 entrants in 1980, the numbers increased to 200 in 1991 and stayed high, at 186, in 1994. Foreign-firms' weight in the UK retail sector has been relatively confined until recently. Only seven of the 50 largest retailers in the UK are foreign-owned. Against the backdrop of this relatively low presence of foreign-owned firms, however, food retailing stands out as an area where their weight, especially following Wal-Mart's acquisition of Asda, is significantly higher. In addition to Asda, Aldi, Netto and Lidl claim small but consistent market shares at the low-cost end of the food retail market. UK's experience with foreign-owned food retailers reflects the general pattern for the world's largest retailers, including those that primarily engage in food retailing.

According to the theory of the multinational enterprise that we detail further below, it is just as important to note the tendencies for internationalization by UK-based retailers. In 2000/2001, of the largest 10 domestic UK retailers, eight of them had overseas operations, but with the exception of Kingfisher, which generated 39.8% of its total retail sales from outside the UK, the percent of total retail sales overseas was below 20% for the rest of the group. Five out of these top 10 UK retailers were food retailers, including the top four. The highest proportion of overseas sales was registered by Sainsbury's, at slightly above 15% (Burt and Sparks 2003: 27) By 2005, however, the share of overseas sales for the leading UK retailer Tesco had risen to 20%. While this figure at first glance appears low compared to the share of sales and revenues among the world’s most internationalized multinational corporations as a whole, it needs to be interpreted within the context of retailing, where even the largest retailers of the world generally rank low in terms of the trans-nationality index, a composite measure indicating percentage of foreign assets, sales, and employment. Walmart, which generates only about 14% of its sales outside of the United States, has a trans-nationality score of only 25. There is reason to think, therefore, that overseas activities will become increasingly important for the retailers based in the mature UK market. Tesco is already among the most internationalised retailers in the world with its current overseas sales shares.

We argue that “multi-nationality”, that is, retailers’ involvement in overseas activities, is an important factor to consider in measuring the productivity of firms in the sector, and hence of the sector as a whole. We base this view on the theory of multinationals (Caves 1996, Dunning 2000), which posits that MNEs’, including their foreign affiliates, are superior in their productivity performance relative to those of purely domestic firms. As Harris and Robinson (2003) point out these arguments are valid only in the long run. In the short run, foreign affiliates might under-perform relative to local firms due to the lack of knowledge of the host market and of local business networks involving suppliers and customers or to cultural differences between the home and host countries. MNEs possess some firm-specific advantage that compensates for the higher costs induced by operating in a foreign market. These advantages may be in the form of superior management expertise, technological capabilities or employees’ technical knowledge. This alleged superior performance of MNEs’ subsidiaries compared with domestic-owned firms has been widely documented in empirical research. For the UK, Griffith et al. (2003) found that foreign-owned MNEs are on average 25 per cent more productive in terms of value added per worker and invest more per worker than both British domestic firms. Interestingly, the same was also true of their productivity as compared with British MNEs and in both the manufacturing and service sectors. Criscuolo
and Martin (2005) show that foreign-owned MNEs in the UK are roughly twice as productive (output per worker) as domestic firms. They also found evidence that UK MNEs are less productive than US owned plants, but as productive as other foreign owned plants.

As the economic rationale for the special treatment of foreign MNEs, policy-makers cite positive externalities generated by Foreign Direct Investment (FDI), with benefits arriving from three basic sources. Firstly, the presence of foreign affiliates increases the competitive pressure on domestic firms in the host country (Barrell and Pain, 1997). Secondly, there are the direct effects of increasing the demand for labour and from the injection of capital. Finally, there are the potential indirect benefits that spill over from foreign to domestic firms. Such “spill-over’s”, both in terms of transfers of technology (Keller 2004), and in terms of upgrading skills in the local labour market as workers transfer between firms (Driffield and Taylor, 2001), clearly can benefit domestic firms, particularly in clustered industries (see Cantwell et. al. 2001). Empirical studies using aggregated and disaggregated UK data have found positive impacts associated with intra-industry, inter-industry and spatial agglomeration effects (Girma and Wakelin 2002, Harris and Robison 2003). However the vast majority of studies investigating the productivity of foreign MNEs in Britain, including these studies, have focused on the manufacturing sector. Hence it is important to see if their conclusions about the impact of FDI hold also in the retail sector. This investigation, in turn, should look most closely at the “ownership advantages” of foreign firms, among which we believe capacities and forms of knowledge transfer require special attention.

Knowledge Transfers Within and Between Firms and Productivity

Knowledge transfer takes place both inside and between firms. It is important therefore to begin with a clarification of the forms of knowledge transfer in the service sector in general, and retailing in particular. Although the role of knowledge is implicitly recognized in the residual “total factor productivity” element in quantitative analyses, there has been surprisingly little detailed research on the contribution of knowledge, in its entirety, as a driver of productivity changes. This partly reflects a failure to theorize these links adequately, and partly the limitations of secondary data, or even of questionnaire-based primary data. Most research therefore has focused on technology, and in retailing specifically on ICT, as a critical feature of knowledge and innovation. However, this provides only a partial understanding of the role of knowledge in productivity changes.

Knowledge in the retail sector has several distinctive features that mediate its relationships with productivity. Den Hertog’s (2002) four-dimensional model of service innovation is useful in highlighting the key areas of knowledge creation and transfer in the retail sector (Wrigley et al 2005): new service concepts (e.g., new store formats into financial services), new client interfaces (e.g. on line shopping), new service delivery systems (e.g. home delivery) and technological options (e.g. client profiling and data mining). These can all contribute to the productivity of individual firms, although in complex ways. However, two particular challenges face retail firms in the utilization of knowledge. These are: protection of proprietary knowledge, and the significance of knowledge of localized consumer markets.

First, in terms of proprietary knowledge, the problem faced by retail firms is that many of these areas of product innovation are highly visible to competitions and so are difficult to
protect (Currah and Wrigley 2004: 9). Therefore, their back-stage process-based knowledge (for example, the application of IT systems to data mining) is a particularly important source of competitive advantage, and of productivity enhancement. Goldman (2001: 223) sums this up in his distinction between the retail offering and the know-how components of retailing. Of course, know-how (in which tacit knowledge is prominent) is also important in all aspects of innovation – for example, a new format may be easily observable to competitors, but this does not necessarily mean that they know how to produce this efficiently and effectively.

Secondly, although in common with most other sectors, the distribution of some types of knowledge across the company is an important objective in retailing, there is also a high degree of sensitivity to the specific characteristics of local consumers. Firms need knowledge of the distinctive multiple markets that they operate in. In other words, they need to use the “embeddedness of stores in local cultures and geographies of consumption within their processes of learning and adaptation” (Currah and Wrigley 2004:6). Traditionally this was seen as stressing the importance of localized tacit knowledge and the need for central management to find ways to harvest and redistribute this within the company. However, recent changes in technology have allowed significant codification of such knowledge, giving the centre greater access to (some forms of) localized knowledge.

In addition to these distinctive challenges of knowledge utilisation and protection, retailing firms also face some of the more generic challenges faced by all firms in knowledge transfer and management. These are discussed below in relation to tacit versus codified knowledge, and inter- versus intra-firm knowledge transfers.

**Tacit versus codified knowledge**

Looking beyond the classic division between tacit and codified knowledge (Polanyi 1966), there are many different forms of the former: in particular, embedded knowledge, which in this case relates to intra-organizational institutions, while en-cultured knowledge highlights the importance of socially situated or place specific knowledge, for example of culturally specific customer behaviour (Blackler 2002). Tacit knowledge is held to be easier to protect than codified knowledge and is therefore particularly important in the competitiveness of individual firms – which, as noted earlier, is particularly important in retailing. Tacit knowledge is articulate in complex ways within companies (Cavusgil et al 2003). At the individual level it can be found in particular skills as well as within forms of abstract knowledge, whilst collective tacit knowledge ‘typically resides in top management’ (p9). The obstacles to knowledge transfers within companies, especially given the dispersed nature of knowledge within organizations, can be formidable (for example Foss and Pedersen 2002) and moreover firms face the challenge of managing intra-, extra- and inter-firm knowledge flows (Nohria and Ghoshal 1997).

The key question for firms is not whether to focus on tacit knowledge or codified knowledge but how to combine these most effectively. In retailing, recent technological developments have rapidly and significantly reconfigured the optimum combination of different types of knowledge, not least because retailing has been one of the biggest contributors to ICT capital deepening over the 1990s in the UK (Oulton and Srinivasan 2004). There are also ownership differences: Bloom et al (2005) have shown that multinationals have higher levels of IT investment than domestic companies, whilst US owned establishments have significant advantages over all other MNCs in this respect. Investment in ICT contributes to timeliness, quality, variety and customisation (Brynjolfsson and Yang 1996), which are critically important in applying knowledge to retailing operations. However, Bloom et al’s (2005) most
striking finding is that 80% of the difference between US and non-US multinationals comes from getting the most out of IT, rather than the amount of IT owned, so that tacit knowledge remains significant.

There are a number of well-known examples of how IT has uplifted company performance. Currah and Wrigley (2004: 130) report on the use of the intranet to generated and distribute target benchmarks and performance ratings for different branches and sectors within a company. Firms are also using IT to customize their operations to local market conditions. It allows them to analyze enormous amounts of data from their own sales information and loyalty card programmes, as well as secondary data and specially commissioned surveys. This leads to greater centralization and, to some extent, to the codification of what had been tacit knowledge about local markets (Wood 2002). For example, the US electronics retailer Best Buy has evolved a ‘customer centricity’ model, whereby stores target highly segmented localized markets (FT 31 October 2005).

**Inter versus intra firm knowledge transfers**

As emphasised earlier, firms potentially have access to intra-, extra- and inter-firm knowledge flows (Nohria and Ghoshal 1997). Of course, these are not discrete options and, as Currah and Wrigley (2004: 2) argue, the key to retail firm performance (and hence productivity) is ‘the interplay between extra-firm (or firm-place) networks of store-based learning and intra-firm networks of knowledge exchange and organizational adaptation.

The choice of transferring knowledge within rather than between firms is the internalisation decision, and is addressed by the transactions cost literature (Coase 1937; Williamson 1985; Buckley and Casson 1976). Barriers to knowledge transfer (which we return to later in terms of ‘absorptive capacity’) create transaction costs which, if greater in external markets (inter-firm transfer) than in internal markets (intra-firm transfer,) promote the growth of establishments under common ownership. Therefore we find multi-plant (and multi-national) firms are associated with production based on knowledge for which external markets are imperfect, notably proprietary knowledge. This applies even to retailing because, as noted earlier, only some forms of proprietary knowledge are easily accessible to competing firms.

However, the decision as to whether to acquire knowledge internally or externally is more complex than this, and is also influenced by non-economic costs related to inter-personal relationships. Insiders may be preferred sources of knowledge for reasons associated with social identities or ease of access, while outsiders my be preferred because external knowledge is more highly valued for status and scarcity reasons (Menon and Pfeffer 2003: 497).

Multinational companies have been more researched than domestic firms in relation to knowledge transfers. Foreign owned firms are, by definition, multinational, and therefore likely to have knowledge intensive operations. It follows that this knowledge is transferred within the firm to foreign affiliates.

The foreign owned outlet can therefore be regarded as an ‘enclave’ of foreign industry within the host economy, providing a unique opportunity to determine the role of knowledge in generating productivity differences between foreign and locally-owned industry. In retailing, it is also important that localised (encultured) knowledge flows from the affiliate to the centre,
given the nature of consumer markets. We return to consider knowledge transfers within MNCs later on this paper, but first we consider inter-company transfers.

**Inter-firm knowledge transfers**

Most recent scholarship has recognized the centrality of different types of knowledge transfer channels and mechanisms. The new approach benefits from a “learning network” of people (Tempest 1999) that is not necessarily related to the formal organisational system of the organisation, nor indeed to the formal boundaries of the firm, as in the example of communities of practice (Wenger 1998). This social constructionist view redraws the map of the firm’s knowledge creation and transfer activities (Davenport et al 1998; Ruggles 1998). There has been a particularly intense debate, notably in economic sociology and economic geography, regarding the milieu that mediate such knowledge transfers. Here we focus on four main competing approaches (but for different perspective see Gertler 2003; and Coe and Bunnell 2003).

- Communities of practice (Wenger 1998). These are informal professional networks where individuals are bound together by shared understandings. They effect knowledge transfers within but also across firm boundaries. They transcend localized boundaries, but trust and shared understanding are enhanced by face-to-face contacts.

- Geographical clusters have high rates of innovation and creativity due to knowledge spillovers as well as shared sets of values and trust, generated through formal and informal networks. Proximity reduces the costs to knowledge sharing in such localized networks, and is especially significant for tacit knowledge transfers. This theoretical perspective is exemplified by the learning regions literature (Maskell and Malmberg 1999). This does not necessarily mean that clusters are enclaves, isolated from external inflows of knowledge. Indeed, external knowledge is also important in such clusters (Bathtet et al 2004), including a steady stream of inward migrant labour as is evident in the success of the Silicon Valley cluster (Saxenian, 1999). This is also a feature of ‘knowledge communities’ (for example, Henry and Pinch 2000) where rapid turnover of staff locally and external transfers of workers are important, alongside cross-cutting supply networks, rapid births and deaths of firms, and the importance of being tuned in to localized gossip and shared discourses, in knowledge transfers.

- Labour mobility is an important source of knowledge spillovers (Arrow 1962), irrespective of whether clusters are involved. For example, Almeida and Kogut (1999) have demonstrated how ideas were spread in the semi-conductor industry through the mobility of key engineers. While labour mobility is managed within multinational companies as a way of effecting knowledge transfers (eg of corporate culture), inter-firm mobility is a potential source of knowledge gain and loss for competing firms.

- Knowledge brokers. Wenger (2000) identified a number of different types of brokering, that is individuals who play a key role in mediating knowledge transfers. ‘Boundary spanners’ (Tushman and Scanlan, 1981), take care of one specific boundary over time. ‘Roamers’ travel from place to place, creating connections, and creating or transferring knowledge. Finally, ‘outposts’ bring back news from the forefront, while exploring new knowledge terrains. Shaw and Alexander (2006) have
demonstrated, in context of retailing, that individuals who held inter-locking directorships of companies played a key role in knowledge transfers relating to supermarkets, in the 1950s.

We still know surprisingly little about the contribution of these different types of milieu of knowledge transfer in any sector, let alone retailing. Currah and Wrigley et al (2004) emphasise the importance of combining different forms of knowledge transfer in context of the combination of top-down and bottom-up processes, and provide some interesting examples, but do not provide a systematic empirical analysis of the subject. However, Alexander et al (2005) do provide a more systematic analysis albeit in context of the post-war growth of supermarkets. They found that the key sources of knowledge transfer were the trade press, government reports, learning from interactions especially from suppliers of equipment, and key individual knowledge brokers. There was also a community of practice based around the Self Service Development Association. Not surprisingly, they found that both tacit and codified knowledge were important, as were different milieu.

**Intra-firm knowledge transfers**

Intra-firm knowledge transfers are also a potential source of productivity changes. ‘Firms are distributed knowledge systems, which means that they are composed of knowledge embodied in individuals and their social interactions’ (Un and Cuervo-Cazura 2004: 28), and the key question is how best to facilitate such interactions. Firms may adopt organization level strategies to facilitate this, seeking to impose a top down structure to capture and redistribute such knowledge (e.g. Brown and Eisenhardt, 1997). Alternatively, epistemic knowledge groups may be created which bring together individuals for specific projects (Thrift 1996; Grabher 2001). In addition, substantial knowledge transactions are informal, involving face to face contacts between pairs or small groups of workers; firms may seek to enhance these through promoting a corporate culture which stresses openness and co-operation. Even where companies are able to capture such knowledge, effective redistribution can be problematic in the retail sector. As Currah and Wrigley (2004: 11) argue “Given that the geography of learning in the retail TNC is inherently store-based, and so fractured … across a mosaic of heterogeneous places, it is … uncertain whether best practices in the retail TNCs can always successfully be adopted and implemented outside the socio-cultural context – the ‘situated knowledge web’ – in which they were originally conceived”

Research on intra-company knowledge transfers has particularly focused on multinational companies. The particular focus on the ownership of activity relates directly to issues of the significance of knowledge in the provision of services and of knowledge transfers both within and between organisations. Productivity differences may arise between firms of different ownership in the same locations, and between different locations (within the UK) for the same (multiplant) firms. Innovation and knowledge transfer, nationally or internationally and either within or between firms, is crucial to productivity (ESRC, 2004),

In the international dimension, the model for the re-use of existing knowledge within the MNE dates from the earliest industrial organisation writings on the MNE (Hymer 1960, 1976, Kindleberger 1969). This portrays the foreign investor as enjoying an absolute ownership-specific advantage over host country firms. Production is organised vertically, with knowledge creation concentrated in the home country and horizontal diversification in the hosts. When the firm’s business strategy is primarily to replicate existing products in new markets or production locations, re-use is optimal (Hansen, Nohria and Tierney 1999). Thus a
codification strategy is ideal, whereby knowledge is recorded for access (eg databases) by the recipients, as part of a “people-to-documents” knowledge management strategy.

However, if the firm aims to develop new and more differentiated products, whatever the market or location, then a strategy for knowledge creation is needed, and the personalisation of knowledge is critical (McDermott 1999). This is described as a “people-to-people” approach, and is reliant on human interaction and tacit knowledge transfer. Where knowledge transfer will take place, and whether it will be contained within the firm or involve external networks, depends on the particular context. Kogut and Zander (1993) have found empirical evidence that the less codifiable and the harder a certain technology is to teach, for example, the more likely its transfer will be to wholly owned operations of a multinational firm. They conclude therefore that a multinational firm's choice of transfer mode is determined by the efficiency of the corporation in transferring knowledge relative to other firms, not relative to an abstract market transaction. It is the multinational corporation's specialization in the transfer and recombination of knowledge that the evolutionary theory of multinationals rests.

The recent emphasis on the possibilities and modes of knowledge transfer using networks external to the firm have also prompted the further investigation of specific processes of knowledge transfer inside the multinational corporation. Inkpen and Dinur (1998), in their longitudinal study of North-American based joint ventures between North American and Japanese firms, identify technology sharing, inter-organizational interaction, personnel transfers, and strategic integration as the four fundamental knowledge transfer processes inside the firm. Gupta and Govindarajan (2000) recognize the significance of the “existence and richness of transmission channels”, but their empirical findings underscore also the significance of “motivational disposition to acquire knowledge” and “absorptive capacity” for the extent of (successful) knowledge inflows into subsidiaries, and of “motivational disposition to share knowledge” and “value of knowledge stock” for the extent of outflows from the subsidiary. Schlegelmilch and Chini (2003), in their more specific discussion of the marketing functions of multinationals, build a model of knowledge transfer where knowledge transfer effectiveness is determined by the a-) development of knowledge transfer capabilities, which are inclusive of channels, infrastructure and processes, b-) organizational distance, and c-) cultural distance.

Different sectors will display different levels of international integration of operations, and depending on whether the products and services produced by a multinational are standardized or segmented within the organisation's network. The mode and degree of international integration of operations in turn relies closely on the motivations of trans-nationalization in a sector or for a specific firm. Peter Dicken has identified the motivations for the transnationalization of a retail firm's operations as including a-) the saturation of the domestic market, b-) intensification of competition in the domestic market, c-) regulatory constraints in the domestic market, d-) perception of profitable opportunities overseas, such as those in some fast-growing developing country markets, e-) desire to exploit a firm's specific advantages in new markets (Dicken 2003: 500). Wrigley argues that in retailing major corporations have specific incentives to trans-nationalize operations because they derive competitive advantages from a-) innovative retail formats, b-) logistics and distribution systems, particularly those that economize on inventory and distribution costs, c-) IT systems and supply chain management, d-) access to low-cost capital for expansion, e-) transfer of “best practice” knowledge, f-) depth of human/management capital resources giving access to a wide range of international management experience, and g-) the ability to source supplies globally (Wrigley 2000: 306-308).
Wrigley further argues that this last ability of global sourcing has not been as important in the operational success of global food retailers as has sometimes been suggested. The non-food consumer product categories in the hypermarkets (that constitute the elite group of retail TNCs) offered the greatest scope to leverage global purchasing scale—the potential in food products was more limited due to differences in national tastes and preferences, perishability issues, and the more restricted overlap between countries in the food products stocked by retailers than might first appear (Wrigley 2000:501).

**Knowledge flows and absorptive capacity**

While this review has emphasised the potential contribution of knowledge transfers to firm performance and hence to productivity, this is contingent on the firm’s absorptive capacity, or its ability to absorb and use knowledge (Bessant et al 2005: 33). Or, as Cohen and Levinthal (1990: 128) put it this is ‘an ability to recognize the value of new, external knowledge, assimilate it and apply it to commercial ends’. Zahra and George (2002) have further refined the concept, arguing that absorptive capacity consists of four distinct capabilities:

- Acquisition – search for new knowledge;
- Assimilation – understanding new knowledge;
- Transformation – seeing how new knowledge can be used in context of the firm’s issues and existing knowledge; and
- Application – implementation of actions enabled by the new knowledge.

Firms posses very different absorptive capacities. The preceding arguments have emphasised that multinational companies have comparative advantages in respect of their absorptive capacity. However, even within retailing MNCs the absorptive capacity may be compromised by intra-branch conflicts, leading, for example, to knowledge hoarding (Currach and Wrigley 2004: 12).

**Skills and Productivity in Retailing**

The third and final section of our review aims to underscore the importance of understanding the relationship between skills and retailing not only through econometric analysis, but, concurrently through an alternate framework where the very notion of skills in the context of service exchange is problematized. Such a double-pronged approach should reveal whether the econometric relationship borne out by quantitative data is supported or contradicted by more process-focused analysis. Furthermore, a critical approach to the notion of “skill” in the service sector also promises to reveal if the productivity-enhancing use of human resources in firms results in other externalities about the quality of the national workforce in general.

This is the second way in which we open up the concept of “knowledge” and its role in retailing. Retailing has certain peculiarities - the intangible nature of output, participation of customers in the production process, and high degree of simultaneity between production and consumption, and different scopes for protecting knowledge in front and back region operations (Goldman 2001) – which exacerbate the significance of knowledge in operations. The following aspects of retail work render knowledge particularly important:

1. the nature of innovative activity (less formal R&D, a central role for information and
communication technologies. In Germany, for instance, 87% of innovating companies in the retail trade sector consider information technology to be important for innovations (European Comisión 2000, p.50), and difficulties of intellectual property protection;

2. the degree of consumers’ knowledge required in delivering certain (especially knowledge-intensive) types of services;

3. sources of consumer information about services; and

4. employee profiles (communication and interpersonal skills) (Cowan, Soete and Tchervonnaya, 2001).

We have discussed the first of these ways in which knowledge is central to retail operations in the previous section, and the second and third ways are of increasing interest in consumption and marketing studies. Keeping our review confined to the supply side of the retail exchange, we now underscore the significance of the final item on this list. Labour productivity is, as aforementioned, central to Total Factor Productivity comparisons. The quality and skills repertoire of the workforce of a retailer will be informed both by the specific location of activity, but then also by the management of labour within the confines of the organisation. A major critical factor in the production and service provision processes of retailers is hence the agency and quality of people involved.

While this account of the significance of knowledge in the work of retailers may appear straightforward, such emphasis on “knowledge” can be misleading due to a common conceptual leap in the treatment of the term in much of the literature. The existing knowledge management literature conflates “knowledge” in the systems and processes of firms and “knowledge capture” activities in which organisations attempt to record, systematise and codify procedures, practices and histories (see, for example Leadbeater 2001, Scarbrough and Swan 2001) with knowledge possessed by workers (McKinlay 2000, Nonaka and Takeuchi 1995, Karremans and Alvesson 2004). This is a serious confusion which obscures (among other things) the nature of knowledge work, the form taken by the labour market, our proximity to government hopes for a “knowledge society” and the type of labour force in currently demanded by employers.

Both knowledge in processes and knowledge in people may add to firm productivity and competitiveness but one does not necessarily involve the other. Indeed, detailing organisational procedures can and does actively deskill work (Davenport and Klahr 1998), while research into the elite of knowledge workers reveals unusual, light touch or widely flouted control mechanisms (Alvesson and Sveningsson 2003, McKinlay 2000, Randle and Rainnie 1997, Robertson, Scarbrough and Swan 2003).

In the UK, retail work is dominated by longer opening hours and the pursuit of “hyperflexibility”, leading to what has been described as “tantamount to a personnel strategy based on zero competence”, zero qualifications, zero training and zero career (Gadrey 2000:26). At the same time, some types of retailing may require detailed technical and specialist knowledge (Darr 2002) and individual retailers may design work so that employees are experts in the products they sell (Gadrey et al. 2001). Although the level of technical knowledge required may vary dramatically between nations (McGauran 2000) in Britain the qualities most frequently sought are soft skills such as communication and customer-focus (Hillage et al. 2002; Nickson et al 2001).
International differences in the skill demands on retail workers can be particularly revealing. In France both retailers and customers expect high levels of technical and product knowledge from salespeople, to the extent that customers will ignore younger workers in favour of their more expert and experienced colleagues (McGauran 2000, 2001) while in Germany rigorous three-year apprenticeships mean that workers are skilled in most aspects of store work and may be flexibly deployed. Ironically, the Anglo-American form of cheap labour and numerical flexibility may be changing German staffing practices (Kirsch et al. 2000). In Britain, by contrast, retail hiring practices may be dominated by judgements about the applicant’s availability for work, proximity to the workplace and gender, something which later feeds in to the way that labour is used and controlled.

The increasing concentration of employment in large companies is another factor that exacerbates these trends introduced by the increased use of ICT in distribution services in general. This may “professionalise” retail work but it may also deskil as central direction on store design, work schedules etc. replaces local autonomy (Sparks 2000a, 2000b, Akehurst and Alexander 1995, Broadbridge 2002). Computers are associated with rising skill levels and there is an identifiable wage premium available for individuals working with them (Haskel 1999, Haskel and Heden 1999, Krueger 1993), perhaps because introducing technology changes the way people work. As a baker, in a piece of research by Smith and Hayton (1999:265) commented, “We don’t just lob a $160,000 piece of equipment on the floor and walk away.” However, the link between technology and up-skilling remains a tenuous one. Certainly from the late 1980s/1990s to date working with PCs and other basic ICT on the job have ceased to be a mark of either skill or status. In other words, technology does not automatically up-skill. Similar wage premia to those available for people working with computers can be found for using a calculator, a telephone, a pen and even for sitting down (DiNardo and Pische 1997; Rubery and Grimshaw 2001). It may be, as Machin (2001) suggests, that computers are now so widespread in the workplace that simply counting them is not a meaningful measure of skill, particularly since not all computer use is complex (Tijdens and Steijn 2005). Felstead, Gallie and Green (2002) reinforce this, pointing out that 77 per cent of people who work with computers use them for only simple or moderate tasks such as printing out an invoice, or using word processing, spreadsheets or e-mail. Only 6 per cent of workers are involved in “advanced” computer work (ibid:60).

The adoption of retail technology has direct consequences for the level of knowledge required by workers, empirical evidence suggesting that this level may vary significantly between different groups of workers. This holds true for the performance of different jobs within the same retail field, firm, or even the same workplace. Wong and Hendry (1999:475-76) identify four major groups of employees in the retail firm: store managers, employees working in inventory and buying, employees who handle retail sales technology itself, and sales staff. They observe that the increased use of ICT in retailing inform the job content of these different groups of workers, as well as their relationship to the organisation’s headquarters, in different ways. While store managers and sales staff are deskilled through the implementation of electronic sales technology, the other groups of employees may now be becoming more knowledge-able.

Another fundamental feature of employment in the retail workplace is the employers’ extensive reliance on more than the technical skills of employees. The issue of “quality labour” is not a straightforward one since what employers are harnessing is not the mechanistic aggregation of an individual worker’s skills and abilities but their capacity to do work (Block 1990). Such a capacity is reciprocal and relational, particularly when
considering the soft, social and customer facing skills most in demand by retailers. As Lafer (2004:117 - 118) points out: “traits such as discipline, loyalty and punctuality are not “skills” that one either possesses or lacks; they are measures of commitment that one chooses to give or withhold based on the conditions of work offered.”

Finally, a significant component of the management of labour in retailing pertains to the management of working time. Despite the focus on customer service, knowledge and commitment, generally associated with full-time, stable employment, 58% of retail employment is part-time (Burt and Sparks 2003), and is characterised by considerable flexibility (Arrowsmith and Sisson 1999). Some stores deploy sophisticated HR management techniques such as psychometric tests (Freathy and Sparks 2000) and merit based pay but these are set against generally low wage rates, rigid control mechanisms and limited discretion (Arrowsmith and Sisson 1999, Broadbridge 2002, Burt and Sparks 2003).

**Conclusion**

In this article we have reviewed the existing literature on productivity in the retail sector and have provided conceptual clarification in a number of key areas. We have established the growing significance of retailing in the UK and made the case that any analysis of the productivity and productivity growth in the country’s economy, as well as the purported productivity gap between the UK and comparable market economies needs to incorporate retailing activity as a core component. Our review has focused on both the factor analysis of productivity outcomes and of specific processes that we deem critical for bringing about these outcomes. As such, it provides the basis for a framework within which future research agendas can be formulated. Research on the productivity of UK retailing is both timely and highly pertinent to a more nuanced understanding of productivity of the larger economy.

We have identified three key rubrics in which such future research would be particularly illuminating. First, we have discussed the measurement issues as they relate to the assessment of productivity in the sector adopting econometric analysis methods. We have specified the complications in the measurement of retail outputs and inputs which require particular attention in the application of this approach. We argue that econometric analysis of retail productivity can be enhanced by focusing on “Total Factor Productivity”, which has five major determinants. These determinants are, in turn, competition and composition effect, planning regulation, the role of Information Communication Technologies, the composition of employment and the role of skills, and the role of multinational enterprises.

While we have provided the framework for a more multi-faceted approach to the capture of productivity levels in quantitative outcomes, we have also suggested that two key areas that have direct and significant impact on these outcomes prompt further qualitative investigation so as to allow for a better understanding of constitutive processes. First, we contend that knowledge transfer and its relation to productivity should be a major topic of investigation. In this, knowledge transfer both within and between firms needs to be considered. Secondly, the analysis of the role of skills in defining productivity outcomes included in the econometric approach needs to be complemented by qualitative research methods that will allow for an understanding of how the knowledge embodied in the retailing workforce is defined, distributed, and experienced in real workplaces.

Our multi-layered and multi-method prescription for the study of productivity in UK retailing
calls forth the utilization of the existing statistical data, as well as the collection of primary data, especially at the firm level. For the processual aspects of knowledge transfer and the use of skills in the retail workplace, sectoral surveys and in-depth case studies emerge as appropriate modes of research. Against the backdrop of the tendencies of internationalisation in the retail sector and based on the theory of multinational enterprises we have outlined above, a comparison of foreign MNEs and British retailers in the UK is one specific methodological tool that promises to be particularly illuminating.

References


Almeida, P. and Kogut, B. (1999), Localization of knowledge and the mobility of engineers in regional networks, Management Science 45(7): 905-917


Cavusgil, S. T., Calantone, R. J. and Zhao, Y. (2003), Tacit knowledge transfer and firm innovation capability, Journal of Business and Industrial Marketing 18: 6-21


ESRC. (2004). The UK's productivity gap: what research tells us and what we need to find out, ESRC Seminar Series.


O’Mahony, M and De Boer, W (2002), Britain’s relative productivity performance: updates to 1999 (London: NESR)


Sparks, L. (2000a) ‘The rise of the large format food store’ in Baret, C., Lehndorff, S. and Sparks, L. (eds.) Flexible Working in Food Retailing: a comparison between France, Germany, the UK and Japan, Routledge: London.
Sparks, L. (2000b) ‘Employment in food retailing’ in Baret, C., Lehndorff, S. and Sparks, L. (eds.) Flexible Working in Food Retailing: a comparison between France, Germany, the UK and Japan Routledge: London and New York


KNOWLEDGE LEAKAGE

Andrew Grantham, Diane Mynors, Kathryn Walsh, Lyn Dodds, Paul Chan, Raphael Kaplinsky

Introduction

In the world of commerce, business’ ability to make knowledge work harder to improve both productivity and competitiveness is ‘the next frontier’ (Drucker 1998; Commission-of-the-European-Communities 2000). This is encapsulated in concepts such as ‘Learning Laboratories’ (Leonard-Barton 1992), continuous improvement (Bessant and Caffyn 1997; Caffyn and Grantham 2003), lean manufacturing (Lamming 1993; Cusumano 1994; Womack and Jones 1996), etc. However, the application of knowledge to the work process is not straightforward. Questions such as ownership of knowledge, knowledge transfer and increasingly complex organisational boundaries make operationalisation difficult. Despite much hype surrounding the role of knowledge in boosting the performance and competitiveness of the economy, several commentators have expressed dissatisfaction with a lack of understanding (and level of abstraction) of the nature of knowledge. Kelloway and Barling (Kelloway and Barling 2000), for instance, commented, “as yet there is little consensus as to what constitutes ‘knowledge work’, making it difficult if not impossible to achieve” (p. 287). They added that the literature around the concept of knowledge has thus far repackaged “old wine in new bottles” (p. 288).

Knowledge is, at the same time, inherently leaky. It permeates organisational boundaries – sometimes intended, sometimes not. Often it is left to perish, with its true value to the business or to the wider world (knowledge as a public good) being overlooked. Moreover, people are defensive with regards to the diffusion of knowledge, but increasing uncertainty and labour mobility generates a circulation of knowledge for good or bad. This review seeks first to synthesise the literature on the concept of knowledge, by borrowing from scholars of knowledge management, industrial organisation, value chains, human resource management (HRM) and trust-based relations amongst others. We have focused on papers that provide either seminal conceptual tools or empirical insights or trends. It is argued that there needs to be greater specificity when discussing the concept of knowledge and a deeper, more holistic understanding can be developed through examining the way knowledge leaks within and across organisational boundaries. This review, therefore, sets the backdrop to developmental work on a taxonomy of what we term as ‘knowledge leakage’. We start with brief discussion on knowledge which clarifies our understanding of the concept for our subsequent work on knowledge leakage and taxonomy building.

Knowledge

**Definition of knowledge**

Knowledge has been well discussed in the literature, and it has different meanings depending on the discipline in which it is used. In here, Awad and Ghaizri’s (Awad and Ghaziri 2004) definition is adopted in which knowledge refers to what is gained through experience or study that enables a person to perform a specific task. It is important to distinguish between information and knowledge for the purpose of this research. For Firestone and McElroy (Firestone and McElroy 2005a; Firestone and McElroy 2005b) information is referred to as a
questionable concept, which could or could not deliver true benefits to an organisation, while knowledge is substantiated much further. It is believed that knowledge has been tested and assessed over time and is a more tangible notion, the benefits of which can be easily determined; it can be an existing structure of information (for example DNA instructions, beliefs or claims) that could facilitate the existing system that developed it to adapt. More generally information is seen as meaningful data, whilst knowledge is information which is interpreted and utilisable such as through a methodology, model or belief system (Wiig 1993; Davenport 1997; Davenport and Prusak 1998; Choo, Detlor et al. 2000; Stenmark 2002). There are many classifications in the literature that fits the above definitions and some of these classifications are reviewed below.

Classifications of knowledge

For economists, knowledge manifests itself in two basic forms: embodied and disembodied (Zeller 2000). Embodied knowledge resides in devices, equipment, machinery, and materials, as well as in human beings in the form of ideas, expertise, skills and routines. It is not, therefore codified or even amenable to codification, and hence is vulnerable to loss or neglect. By contrast, disembodied knowledge is accessible to us all through databases, manuals, patents, specifications, IPRs, scientific books and journals. In general, the greater the codifiability of knowledge, the lower the barriers to entry. Codification is thus a danger, nevertheless at the same time, the absence of codifiability may often mean that firms may fail to systematise their knowledge base and maximise the returns from their knowledge flows. Styhre (Styhre 2004)

Knowledge can also be classified as either explicit or tacit knowledge; the former being easily codified and the latter being embedded in the human brain and cannot be expressed easily (Grover and Davenport 2001). The concept of tacit knowledge has previously found fascination among organisational/management theorists and the knowledge management research community and is derived from the philosopher Polanyi (Polanyi 1958). In discussions around tacit and explicit knowledge (Nonaka and Takeuchi 1995) there remains a belief that tacit knowledge can be rendered explicit and hence shared and extensively utilised (Marshall and Sapsed 2000).

Styhre (2004) believes that the demarcation between explicit and tacit knowledge is a false dichotomy and that explicit and tacit knowledge is intertwined; a continuum between intellect (objective knowledge) and intuition (subjective understanding). Moreover, Styhre (2004) argues that knowledge is only useful in a social, contextual and holistic setting and therefore should be examined within the same setting. Styhre’s (2004) review of the literature on tacit knowledge highlighted Boisot’s (Boisot 1998) distinctions of tacit knowledge to include: a) matters that are said because everybody understands them and takes them for granted; b) matters that are not said because nobody fully understands them, and thus they remain elusive and inarticulate; and c) matters that are not said because some people can understand them. Styhre (2004) states that knowledge management theorists have been emphasising distinction C, the present study is addressing all the above-mentioned Boisot’s variants.

Styhre (2004) invokes Popper (Popper 1983) in the proposition of a continuum between intellect and intuition Popper (1983) makes the philosophical suggestion that knowledge revolves around a continuum of three worlds; the world of objective knowledge (world 3), which can be derived from the physical world of objects and states (world 1), but which is only effective through subjective human experience (world 2). Popper’s (1983) perspective of knowledge is not dissimilar to Gibbons et al. (Gibbons, Limoges et al. 1994) Mode I
(scientific knowledge) and Mode II (application oriented) knowledge; or Mukherjee et al.’s. (Mukherjee, Laprè et al. 1998) distinction between conceptual and operational knowledge. The former relates to know-why and the latter relates to know-how. Conceptual learning is the process of acquiring a better understanding of cause-and-effect relationships, i.e. the acquisition of know-why. Operational learning is the process of obtaining validation of action-outcome links, i.e. the acquisition of know-how.

Billett (Billett 1997) identified knowledge as propositional (i.e. Mode I), procedural (i.e. Mode II) and included a third category: dispositional, i.e. learnt values, attitudes and interests that predispose the acquisition and treatment of knowledge. For knowledge to be meaningful, Fleck (Fleck 1997) propose that it needs to be within an appropriate contexts such as: domain (an area of expert focus which provides a particular view; for example the accountant vis-à-vis the engineer); situation (assemblage of people and objects in discourse at the same point in time) and milieux (character of the immediate physical and social environment in which knowing activities take place such as workplaces over time).

However, there is a wider debate around types of knowledge and the above-mentioned classifications of knowledge fits several similar classifications found in the literature. The authors are aware that there is extensive literature on knowledge, but it is beyond the scope of this text to discuss these literatures any further and will focus on the concept of knowledge leakage in the following sections to explore how knowledge flows affects an organisation.

In reviewing these seminal works that define knowledge a research gap is exposed. The level of abstraction renders the concept somewhat opaque at an operations level. Taylor (Taylor 2002), for instance, lamented on the level of abstraction where knowledge is concerned, and suggested that “[h]onest probing is needed now, rather than glib answers”. In order to operationalise knowledge effectively, practitioners seek guidance on knowledge impacts on the firm and a toolbox of indicators by which to leverage it for competitive advantage. A conceptual and empirical investigation of knowledge leakage, consequently, lends itself to this task. The next section provides some insights into the nature and extent of knowledge leakage from firms.

**The concept of knowledge leakage**

There are different terms used in the literature to refer to the concept of knowledge leakage. Terms mentioned include knowledge seepage (e.g. (DiRomualdo 2004; Kingston 2004; MacDougall and Hurst 2005); knowledge transfer (e.g. (Bhattacharya and Guriev 2004; Huang 2004; Kingston 2004; MacDougall and Hurst 2005; Marti and Fallery 2005); knowledge loss (e.g. (Huang 2004; MacDougall and Hurst 2005); knowledge disclosure (Bhattacharya and Guriev 2004) and knowledge leakage (e.g. (Lamming 1993; Twigg 1997; Bhattacharya and Guriev 2004; Vohringer, Kuosmanen et al. 2004; Annansingh 2005). In the references cited above, the key consideration is the movement of people.

For firms, it is important to be able to assess whether knowledge leakage is occurring and to what extent it is damaging or can be rendered positive. Annansingh, (Annansingh 2005) defines knowledge leakage as “the possibility of information or knowledge that is critical to the organisation being lost or leaked – whether deliberately or unintentionally – to a competitor or unauthorised personnel”. This is perceived as negative knowledge leakage as sole ownership of knowledge, leaks away from the origin it may lead to a loss of competitive advantage. Knowledge leakage can also be positive, Vohringer et al., (Vohringer, Kuosmanen
et al. 2004) defines knowledge leakage in positive terms when it occurs in the form of information spillovers between project partners and projects.

It is essential that organisations differentiate between the types of knowledge that can be leaked and have an impact on the organisation. MacDougall and Hurst (MacDougall and Hurst 2005) and Matusik and Hill (Matusik and Hill 1998) make the distinction between what can be leaked as public knowledge and private knowledge. Public domain knowledge they refer to as: knowledge that resides in the external environment. While private knowledge is referred to as the key competitive knowledge such as an organisation's unique routines, processes, documentation and trade secrets. Matusik and Hill (1998) believe that it is a competitive threat if the private knowledge is leaked into the public domain.

Mansfield’s (Mansfield, 1985) study investigated a number of sectors, including chemicals, pharmaceuticals, petroleum, primary metals, electrical equipment, machinery, transportation equipment, instruments, stone clay and glass, fabricated metal products, food, rubber and paper, and the rate at which information about development decisions leaked out to competitors or into the wider business community. It is reported that for one fifth of firms, leakage occurs within six months. At best, such critical information is in the hands of rivals between 12 and 18 months. There are some sectors – chemicals and glass – in which leakage is slower. Leakage regarding process innovations is slower due to better internal capabilities to generate them with less communication and interaction.

It is important that organisations distinguish between the consequences of knowledge sharing (both parties have access to it) and absolute knowledge loss (that is, one firm’s gain is another firm’s loss, as in the flow of people. The routes though which knowledge leakage occurs are identified as informal communication networks, professional networks and employees moving from one firm to another, which will be discussed in more details in the human factors section below.

What is clear from scholars’ work in defining and studying leakage is how little we know about the value of the leakage. For example, in the Mansfield study, whilst information and/or knowledge leak at an alarming rate, the extent to which it truly impacts on firms’ competitive advantage is neither elucidated nor quantified. It is unclear whether the know-how is leaking or inconsequential crumbs of information. Moreover, its bias towards knowledge loss diverts the focus towards policy and away from operational matters such as productivity. It also diverts practitioners away from assessing their knowledge resources and what they need to succeed. The next section makes the link between knowledge leakage and firm-level attributes such as dynamic capabilities, core competences and supply chain development.

**Implications of knowledge: different perspectives**

**Dynamic capabilities and core competences**

Dynamic capabilities and core competences stem largely from the resource-based view of the firm, informed by the work of Wernerfelt (1984) and Barney (1991) (Wernerfelt 1984; Barney 1991). Teece et al. (Teece, Pisano et al. 1997) argue that “[w]inners in the global marketplace have been firms that can demonstrate timely responsiveness and rapid and flexible product innovation, coupled with the management capability to redeploy internal and external competences”. (p.7) Prahalad and Hamel (1990), on the other hand, talk about core
competences as capabilities that are inimitable by rivals, providing organisations with competitive advantage.

There are three key configurations of capabilities to consider. First there are internal capabilities that are explicit and homogenous such as product development and strategic decision making which pool business, functional and personal expertise (Eisenhardt and Martin 2000). Second, there are internal capabilities that are tacit and heterogeneous such as knowledge resources (Grant 1996; Kogut 1996). Thirdly, there are critical inter-relationship capabilities such as commercial alliances/inter-firm cooperation (Lorenzoni and Lipparini 1999; Eisenhardt and Martin 2000; Schmitz and Knorringa 2000; Bessant, Kaplinsky et al. 2003).

Dynamic capabilities, consequently, are restricted to fostering competitive – but not sustained – advantage for the firm. The possession of dynamic capabilities, therefore, may be seen as a necessary condition for business development – particularly in the realms of balancing the demands of exploitation (fully realising the value of existing platforms and products) and exploration (investing in new platforms and technologies). From a strategy perspective, the concept of core competences is also significant.

Prahalad and Hamel (Prahalad and Hamel 1990) offer three tests for firms seeking to identify their core competences:

- that a core competence provides access to a wide variety of markets. This may be either a technological or a management capability.
- that customers see that it adds significant value to the end product.
- that it is difficult, if not impossible to copy. (p5-6)

Whilst rival firms may be able to acquire technologies that comprise a core competence, this may be restricted if the core competence is a ‘complex harmonization’ of technologies production and managerial routines.

Whilst the criteria for the test may at first appear unremarkable, the inability of firms to identify their core competences has often led to their demise or sub-optimal performance. Prahalad and Hamel note that firms tend to be better at listing their capabilities than their core competences (hence the value of the test). It is also argued that firms are unlikely to build expertise in more than 5 or 6 competences; and to attempt to do so may be strategically unwise, or indeed to resort merely to listing capabilities.

Another way of looking at core competences for firms is to focus on core products to act as ‘lynchpins’. Using a biological analogy, at the base is neither the product nor the business. Instead the limited core competences form the ‘roots’ with the lynchpin products as the stem or trunk. Further up is located the appropriate business unit to deliver the end products (or leaves) for the ultimate customer or consumer. Their most compelling example is Honda’s engines which enable the company to compete in a number of markets and is derived from complementary technological, production and managerial competences.

Also pertinent are rents and barriers to entry. The Ricardian concept of rent describes a situation where the parties who control a particular set of resources are able appropriate rents out of scarcity value – and maintaining it – by erecting barriers to entry to eradicate or limit effective competition.(Kaplinksy 2005 p62). Schumpeter provided an analytical framework to show how scarcity can be constructed (Schumpeter 1961). He distinguished the process of
‘invention’ (having an original idea, a ‘new combination’ in his words) from that of ‘innovation’ (turning a new idea to commercial advantage). Entrepreneurship is defined in the act of innovation. If this innovation is difficult to copy, then the entrepreneur earns a super-profit which exceeds not only the cost of the invention and the associated innovation, but the returns to economic activity in other activities which are less well protected from competition. Over time this innovation is copied (the act of ‘diffusion’) or superseded by a new, superior innovation. It is this ‘Schumpeterian motor’, the search for producer rents, that spurs the innovation process and subsequent diffusion and that drives forward economic growth. For Schumpeter, the entrepreneurial rents were almost always dynamic.

By way of caution, Scarbrough (Scarbrough 1998) attacks the resource-based view of the firm for resulting in a weak link between competencies and performance, as he maintains “little attempt to demonstrate the mechanical links, between competencies and performance, other than in the broad terms of the root and branch metaphor propounded by Prahalad and Hamel”. (p. 224: original emphasis).

Consequently, for Scarbrough (Scarbrough 1998), “theorists attempt only the sketchiest account of the nature of resources and competencies, preferring to identify them inductively from evidence on a firm’s functional outputs or competitive advantage” (Scarbrough 1998, p223).

This resonates with the argument that the concept of organisational knowledge has hitherto remained abstract, as Styhre (Styhre 2004) argues ‘the doctrine of tacit knowledge is based on a belief in a rational human mind that can structure, organise and make sense of complex realities; when this process is not fully understood, the forms of knowledge generated are called tacit knowledge; tacit knowledge is thus an anomaly in a representativistic paradigm, a failure to express what we think we know; it is rare that the assumptions and underlying ideas of the notion of tacit knowledge are articulated or discussed.” (p185’).

Another problem with drawing on these perspectives is consistency in levels of analysis, especially where the impacts of knowledge on organisational performance are concerned. With respect to productivity and competitive, the operations management literature is micro-level in its emphasis (for example, (Mills, Platts et al. 2003). (By contrast, productivity studies with their macro economic objectives avoid discussing the human factors that deliver productivity.) (Crafts and O'Mahony 2001).

This raises some issues with value chain models which conceptually at least occupy macro-level space. It can be used as a tool for businesses to inform a strategy of upgrading; that is, to (re)position themselves in the chain in order to realise as much value from their normal activities as possible. For example, moving up the chain from mining raw materials to processing or sub assembly to final assembly. In the context of knowledge leakage the value chain framework can be used to map knowledge transfers and assess the extent to which these can be, and are, used in the upgrading process. For example, in moving from sub-assembly to Original Equipment Manufacture (OEM). (Porter 1990; Grantham and Kaplinsky 2005).

Knowledge leakage is by definition dynamic and has an impact across the value chain, particularly with respect to inter-firm linkages (with suppliers; and increasingly with customers, competitors and other organisations such as universities and trade associations). Global Production Networks (GPNs) (Ernst and Kim 2002) are essentially a set of inter-firm relationships that bind a group of firms into a larger economic unit (Sturgeon, 2001 p2). They
have a bearing because GPNs are potential conduits of knowledge diffusing between ‘member’ businesses across boundaries. The direction of knowledge transfer is often between businesses in the developed world and those acting as (lower-tier) suppliers in low-cost locations mediated by ICTs (information and communication technologies). Hence it is immensely valuable in the context of economic development; though the consequence of knowledge leakage is less well understood. And the globalisation is driven partly by the search for distant markets with a view to amortizing often huge Research and Development (R&D) costs that cannot be covered by sales in domestic markets alone.

Consequently, rarely do single concepts provide a full picture of firm behaviour in either predictive or prescriptive senses. A series of complementary concepts are useful in expanding understanding in recognising the primacy of inter-firm linkages in the attainment of systemic efficiency.

In Reed and Walsh’s paper on enhancing technological capability through supplier development (Reed and Walsh 2002) they describe a historical view of large companies as being ‘in-house’ designers with funding for internal R&D and spare resource for speculative work, which allowed them to be forward thinking as well as developing current products. The change within manufacturing industry to supply chains development and with the focus shifting to core competencies means the responsibility is shifting for the advancement of technology and smaller suppliers may not have the resources to plan for new technologies (or ‘technology lookahead’ activities), or may be unaware of the needs of the original company (p231). They also suggest that original equipment manufacturers (OEMs) may not share [leak] advanced knowledge of future needs with suppliers unless specific action is required which may have implications for maintaining technological capability. Their study of two aerospace and defence organisations found that their supplier development programmes were found to enhance supplier technological capability. The EU productivity report showed the EU outperformed the US in industry groups where innovations rose from in house R&D (O’Mahory and van Ark 2003 p12) R&D investments linked to increased productivity. Small suppliers are less likely to ‘look sideways’ at potentially disruptive technologies emerging in other industry sectors (Reed and Walsh 2002 p232).

**Agency Theory and Stakeholder Theory**

Shankman (Shankman 1999) maintains that “[m]odern agency theory [bears] roots of which lie in the field of organisational economics (p. 319)”. Drawing on seminal theorists like Coase (Coase 1937) and Williamson (Williamson 1975), agency theorists have focussed the relationship and governance structures between the agent and the principal. This consequently sets the basis for modern-day contractual arrangements. However, agency theory, it is argued, delimits the level of holism in understanding the nature of organisational knowledge as a dynamic concept. It assumes rationalistic behaviour that invokes the debate over the structuration of knowledge raised earlier. By contrast, stakeholder theorists like Freeman (Freeman 1984) are more encompassing as they recognise that corporations are often influenced by a range of stakeholders, both within and across organisational boundaries. Indeed, with stakeholder theoretical approach, “the role of management is seen as achieving a balance between interests of all stakeholders (Shankman 1999 p322)” . Undoubtedly, the study of organisational knowledge requires greater depth of holism. With political trends towards embracing stakeholder approach (e.g. New Labour government’s stakeholder approach in the UK), our study of organisational knowledge leakage within and across boundaries represent a novel, empirical opportunity to achieve such holism.
**Knowledge and Business Improvement Methodologies**

A discussion of business improvement methodologies is useful in that first of all they tend to be focused on the firm or a group of firms, and second, they are often rigorously deployed in businesses with robust metrics used to quantify their efficacy. Moreover, knowledge leakage is often a key element of the successful implementation of business improvement methodologies; for example, Continuous Improvement (Bessant and Caffyn 1997; Caffyn and Grantham 2003). Total Quality Management (TQM) owes much to the collective contributions of Crosby (Crosby 1979), Deming (Deming 1982), Juran and Gryna (Juran and Gryna 1993) and Feigenbaum (Feigenbaum 1991). More critical and dynamic debates about TQM are emerging (McAdam 2004) and Taylor and Pearson, (Taylor and Pearson 1994). From a strict operations perspective, lean production systems (Womack and Jones 1996; Shah and Ward 2003) now prevail and operationalised through improvement methodologies such as Continuous Improvement and supply chain management (Bessant, Kaplinsky et al. 2003).

It is in this domain where the innovation management and industrial relations literatures differ in their analyses. The innovation management literature is at best agnostic about the implications for employees with respect to employment. The industrial relations literature has legitimate concerns about deskilling. For instance, Grugulis and Bevitt (Grugulis and Bevitt 2001) investigated the impact of Investors in People on a hospital trust, and suggested that short-term performance targets and business needs might contravene employees’ training requirements and thereby lead to erosion of skills. In a similar vein, Grugulis et al. (2003), through case study research, found that employers often chase organisational flexibility at the expense of individual employee discretion, resulting in deskilling. More recently, Nyhan et. al. (2004) blamed modern management for “not paying a great deal of attention to ensuring personal learning benefits for employees and workers” and envisaged a repetition of the “reality for many workers, today, is a reincarnation of Taylorism in the form of neo-Taylorism” (p. 69) (see also (Thursfield 2001)) However, national policy interventions such as the promotion of lifelong learning for sustaining a knowledge economy have a strategic outlook (Harrison and Kessels 2004).

Complexity, cost and globalization ensure that firms increasingly interact across the boundaries of (internal) functions (e.g. planning, production), the firm (e.g. industrial clusters), and nations (for example, outsourcing). Increasingly collaboration mediates new product development (Wheelwright and Clark 1992; Oliver and Blakeborough 1998) which is challenging in the context of innovation management; for example, protection of intellectual property (Tang and Molas-Gallart 2005) and effective functioning of dispersed teams (Sapsed, Gann et al. 2005) and utilizing accrued knowledge through Learning Networks (Bessant and Tsekouras 2001).

**Knowledge and knowledge Intensity**

Autio et al define knowledge intensity as the ‘extent to which a firm depends on the knowledge inherent in its activities and outputs as a source of competitive advantage’ (Autio, Sapienza et al. 2000 p913). The measurement of knowledge intensity in firms and the economy as a whole is problematic and often blunt. R&D activity is a widely used indicator captured by R&D expenditure, R&D expenditure, education and training, software, market research, design, patents, licenses, capital investment, intermediate goods acquisition (Smith 2002), R&D/Sales (Lepak, Takeuchi et al. 2003). Others deploy Tobin’s q ratio which measuring the relationship between a company’s market value and its replacement value or its physical assets (higher ratio suggests greater knowledge intensity, see (Sveiby 1997; Swart and Kinnie 2003).
Other measures include: level of education of employees (Smith 2002) and type of workforce employed (knowledge-based, job-based, contract, alliance/partnership etc, see (Lepak, Takeuchi et al. 2003).

There is also a recognised hierarchy of knowledge intensive tasks based on the problem being solved – in recognition of the fact that there are different ways of using knowledge (Shadbolt and Milton 1999). Lepak et al (2003) argue that intensity is also captured by the degree to which organisations rely on standardised knowledge, or on unique employee contributions. Perhaps more interesting is indicators offered by Roper and Crone (Roper and Cronet 2003). They discuss stock of knowledge within an organisation measured through an assessment of managerial and production techniques (such as the use of IT systems). At the subjective level, Autio et al have used Likert scales to assess organisations’ knowledge intensity (Autio, Sapienza et al. 2000).

*The absorption and transfer of knowledge BETWEEN organisations*

We have already discussed an organisation’s need to maintain stocks of competitive, or private knowledge (such as knowledge about core competences). The stocks refer to the accumulated knowledge within the organisation. However in order for the organisation to continue to appropriate rents (or to keep the Schumpeterian motor turning), and to maintain effective barriers to entry, they must also maximise the positive and minimise the negative consequences of the absorption and transfer of knowledge into and out of the organisation. Absorption refers to flows of knowledge INTO the organisation. Transfer refers to flows of knowledge OUT of the organisation.

Many authors note the importance of knowledge sharing among organisations. Efficient data sharing can increase resource utilisation and reduce costs (for example see (Stefansson 2002). Critically, the elementary factor is the development of effective mechanisms for sharing knowledge.

Many companies now understand the significance of proprietary data (data upon which the company is managed) and shared data (shared through a contract or standard to which all parties agree). This has increased in importance as many companies are outsourcing to third parties, (Stefansson 2002). The supplier is typically cast as the recipient of knowledge – though this is not always the case. Roper and Crone (2003) recognize the “mutual interdependence and the value of more effective knowledge coordination than that associated with more adversarial supply-chain relationships” (p340).

Roper and Crone’s work regarding knowledge complementarity was developed to look at the relationship between partner’s knowledge portfolios and the knowledge coordination activity that takes place between them (although no direct relationship was found). One of the components of their work was defining knowledge sharing activities (or knowledge coordination). By defining these activities they sought to increase understanding of the information flows between organisations.
Roper and Crone cite an interesting discussion from Buckley and Carter (1999). They discuss knowledge sharing and more specifically the consequences associated with the sharing of knowledge. They suggest three distinctions:

- additive complementarity – where the sharing of knowledge yields immediate gains;
- sequential complementarity – where the sharing of knowledge stimulates further knowledge seeking behaviour; and
- complex complementarity – where the knowledge possessed by each party is of value to the other, and where reciprocal knowledge transfer [and absorption] occurs (Roper and Crone 2003, p342).

Buckley and Carter also distinguish between incidental knowledge sharing (which they suggest is more widespread) and intentional knowledge sharing (with the specific intention of sharing knowledge (Roper and Crone 2003, p342).

Roper and Crone also suggest two other factors that may influence the way organisations share knowledge, and which may affect the success of the absorption of that knowledge by the recipient organisation.

Firstly is the importance of willingness to take part in knowledge coordination activities. For example Wong (1992) found evidence that knowledge transfers from Singapore based multinationals to their local suppliers were mainly regarding generalised manufacturing practice (know-how) and very rarely regarding design capability. Roper and Crone conclude that knowledge transfers are associated with organizational and production knowledge and not managerial knowledge. A supplier’s willingness to absorb transferred knowledge is also important, and depends on, for example, the perceived profitability of doing so.

Secondly Roper and Crone hypothesise that if a knowledge gap between a firm and supplier is beyond a certain size, the recipient will be unable to assimilate all the transferred knowledge (resulting in knowledge leakage, (Young and Lan 1997). Unfortunately they do not discuss this in relation to their findings.

There are numerous ways in which knowledge is absorbed and transferred between organisations. Consider for example absorption through the purchasing of intermediate or capital goods embodying knowledge, or transfer through the selling of licenses to use
protected knowledge (Smith 2002), or perhaps the exploration of markets prior to internationalisation in order to absorb new knowledge to compete and grow in markets about which the organisation has little prior knowledge (Autio 2000).

The success of these flows of knowledge is dependant of many factors. In this context, Shadbolt and Milton make some interesting comments about the transfer of knowledge from experts:

- Experts vary in how well they can articulate knowledge;
- Experts vary in how well they recall information in a given context;
- Experts vary in their ability to recall the same information in different tasks;
- Experts may vary in the validity of their knowledge and the extent to which they misinterpret information, and are biased or error prone (1999).

They also mention the importance of selecting the correct way of communication knowledge suggesting “a well-chosen analogy, anecdote or diagram can make all the difference when trying to communicate a difficult idea to someone” (p311), and that “ordinary language is the main form of communication, yet it is so full of jargon, assumptions and ambiguities that people often fail to understand what others are trying to say (p313).

Autio et al suggest that the assimilation of new knowledge is less effective the more it requires the unlearning of old knowledge, and that new knowledge is most effectively absorbed when it is in the proximity of existing knowledge (2000) Autio, Sapienza and Almeida also comment that firms that internationalise at a later stage are likely to have developed competencies constraining what they see and how they see it (with reference to new knowledge) which they refer to as “cognitive impediments to learning”. Therefore exposing themselves to a higher risk (p10).

**The absorption and transfer of knowledge WITHIN organisations**

It is not just the absorption and transfer of knowledge between organisations that can effect organisational competativeness, but also how knowledge is shared within an organisation. Meer-Kooistra and Zijlstra specifically mention the ability to replicate knowledge held within the expertise and experiences of people within an organisation as a key mechanism for transforming knowledge into economic value (2001, p456). They explain the knowledge held within people can create value, due to cooperation and coordination by which exchanges of knowledge occur and new knowledge can be developed. They go on to explain how these experiences and expertise can be captured and codified (for example through patents, copyright, databases, etc) which they categorise as intellectual capital (IC) assets and those that are tacit, referred to as IC skills (2001).

Yanow (Yanow 2004) suggests one way that knowledge is lost is through not taking advantage of sources of information, the example he gives is of delivery drivers who are in ‘intimate’ contact with shop owners and customers are not consulted when market information is sought. In this example knowledge is absorbed from outside the organisation, but is not effectively transferred within the organisation. He suggests that this information is not sought because it is devalued because it is held by those at lower levels of the organization. He goes on to explain that these types of knowledge loss usually occur over hierarchical lines, although it might also be found across generations (new employees not learning from old employees, etc).
Und and Cuero-Cazura (Und and Cuero-Cazura 2004) believe that knowledge is created (development of knowledge stocks) during interactions among individuals, and especially through individuals who possess different knowledge sets, through which knowledge is shared and transformed. They also believe that interaction is how the flow of tacit knowledge from one individual to another occurs. They also explain that these interactions rely on:

- the willingness of individuals to share their knowledge, and
- that individuals with different knowledge sets understand one another.

Styhre (2004) also makes a suggestion that knowledge is only useful in a social, contextual and holistic setting and therefore should be examined within the same setting. That by codifying knowledge, some of the knowledge will be lost.

The Bergsonian view Styhre refers to in the title of his paper is taken from Henri Bergson. Styhre draws on his comments and argues that the notion of tacit knowledge is produced in a “rationalistic, foundationalist doctrine of knowledge wherein that which cannot be fully represented” (p178). And that these ambiguous and confusing attempts at representation are subject to misunderstanding or are excluded from the discussion. He also argues that tacit and explicit knowledge are not discrete categories, but always coexist in one another.

Styhre also mentions the importance of ‘intuition’ as well as tacit knowledge, as part of a persons ability to make sense of information (p17)

**The interplay between knowledge and people issues**

From the definitions of knowledge provided in Section 2 above, it can be seen that knowledge, especially of the tacit nature, resides in individuals and it is the understanding of this knowledge that has driven much knowledge management research.

Much research exploring the links between human resource management (HRM) and knowledge focuses on the efficacy of knowledge sharing. For example, Swart and Kinnie (2003) explored through case study research HR practices of recruitment and selection, resource development and participation on knowledge integration within distributed knowledge systems. They emphasised the sharing of knowledge and highlighted “the provision of social supports for interconnecting various stakeholders in the knowledge sharing process (p70); (see also (Treleaven and Sykes 2005). Sharing knowledge about organisational activity including process, changes in products and services among the workforce is also seen as beneficial to organisational efficiency and effectiveness since the act of sharing such knowledge can make employees feel involved in making the organization successful. Indeed, social support mechanisms such as story-telling (Kleiner and Roth 1997), communities of practice (Wenger 2000) and learning ‘laboratories’ where complex organizational ecosystems integrate problem solving, internal knowledge, innovation and experimentation, and external information (Leonard-Barton 1992) have seen increasing adoption in practice by many sectors.

The importance of knowledge sharing amongst the workforce cannot be over-emphasised. Current trends necessitate the consideration of effective knowledge sharing. For example, the problem of aging in Europe could mean 20% of the European population will be over 65 years old by 2025 (Farrell and Knight 2003). Therefore, this means that knowledge of
workers needs to be harnessed effectively so that the impacts of knowledge loss amidst a rapidly retiring working population and a constricting labour market can be mitigated.

More crucially in the short-term, a recent study revealed that 90% of the UK workforce have permanent contracts; yet, these workers stay in the same organisation for merely an average of seven years and four months (Taylor 2002). With such trends, the harnessing of knowledge from the workforce becomes an even more pressing priority.

The literature associating knowledge loss and people issues tended to look at the implications of people movements. For example, DiRomualdo (2004) coins the term “knowledge seepage” when discussing about the shifting of jobs from one set of employees to another in the context of off-shoring. According to DiRomualdo (2004), the conditions of off-shoring are mostly not useful to a complete transfer of knowledge and know-how, particularly when the workers requested to transfer their knowledge face redundancy. In addition, expert staff may also seek jobs elsewhere as they see their colleagues being redundant and therefore feel that everyone is expendable.

Similarly, Treleaven and Sykes (Treleaven and Sykes 2005) state that as a social process knowledge sharing within and across organisational networks support spontaneous practices (which consecutively enhances organisational capability); such organisational knowledge is reported to be lost rapidly where staff turnover is high or where demoralised staff withdraws their organisational loyalty. Knowledge seepage is also reported to happen when all human expertise in an area is steadily lost as the specialists and users become reliant upon the system (Kingston 2004). This particularly is considered as a major risk to organisations mainly in the commercial climate as reorganisation is influential and regular. Thus, DiRomualdo, (2004) suggests that “risk factors” should be identified and the effect of shifts in organisational change can be assessed.

Treleaven and Sykes (2005) also argue that restructuring around a company’s managerialism without consideration of differential appraisal of worth, results in losses of organisational knowledge. For example, when skilled practitioners are retrenched and new management positions are introduced to prioritise financial training and management know-how in company environment, it results in the company losing personal knowledge developed through the application of professional training and experience in specific fields. Treleaven and Sykes (2005) maintain that losing such tacit and heuristic knowledge of workforce (particularly in client services) may not merely influence momentous activities with clients, but also greatly affects the competence of the organisation to give efficient services in various conditions.

Littler and Innes (Littler and Innes 2003) have also examined the implications of downsizing and deskilling. Using a longitudinal Australian dataset, they test a series of propositions relating to the knowledge impacts of structural changes in businesses across sectors. Their ultimate question is ‘does downsizing improve the knowledge-base of organizations?’, to which the answer is an emphatic ‘no’. However, in trying to explain this in the Australian context, they say: ‘[t]his linkage may be the outcome of various mechanisms – intentional strategy or the fact that firms do not adequately measure or monitor skill profiles, or control skill losses.’ (p86). This has distinct implications for firms and their knowledge leakage strategies. Moreover, they find that deskilling is most likely in manufacturing firms when they downsize.
The literature also alluded to the impacts of engaging contingent knowledge workers or temporary workers. MacDougall and Hurst (2005), for instance, state that the deployment of contingent knowledge workers could be an effective way of tapping into public domain knowledge found in the external environment, which in turn could contribute to the optimisation of the firm’s working practices. Nevertheless, Matusik and Hill (1998) cautiously warned that revealing contingent workers to private, key competitive knowledge (e.g. firm’s unique routines, processes, documentation and trade secrets) is risky, as they potentially could threaten to leak this private knowledge back into a public domain.

The importance of worker experience has also been explored in the literature as a proxy for knowledge. In a report discussing the current state of the UK tool making and die industry, Mynors et al. (2004) also mentioned a significant experience-loss mechanism. The report concluded that there is a need for young people in the engineering sector. Many companies depend on the skills of particular key employees who completely retain the knowledge necessary to perform their job. The main worry for these companies is when these key members retire; the knowledge is lost from the company completely. As mentioned at the outset of this section, the risks associated with a greying working population threaten the competitiveness of Europe. In contrast, Leonard-Barton (1992) through participation on an R & D project found that “line manager[s] [need] to have authority over production and projects, and a young engineer is not a good choice to [since they] lack experience in formal problem solving (p1322).”

What we can draw from these conditions is the realisation that (organisational) knowledge – if it exists at all – is a process rather than an artefact. The conduit is people, and whilst it can be deposited in some repository or captured in some software solution, more significant is its diffusion amongst people in its application to problem-solving across functions. As can be seen, the interplay between knowledge and HRM is yet to be fully explored. There is little consensus as to what knowledge really means when discussed in relation to human resource management. Proxies such as skills, experience and organisational routines, amongst others, have in the past been conveniently selected to represent knowledge in the studies cited, often based on the researcher’s particular perspective. Moreover, studies have not necessarily looked into the relationship between knowledge and productivity per se, but rather emphasised the issue of knowledge flows and knowledge sharing for the purpose of performance improvements.

The interplay between knowledge management and HRM is yet to be fully explored. Storey and Quintas (Storey and Quintas 2001) focuses on knowledge sharing as a means to manage knowledge effectively. Swart and Kinnie (2003) who explored through case study research HR practices of recruitment and selection, resource development and participation on knowledge integration within distributed knowledge systems. They emphasised the sharing of knowledge and highlighted “the provision of social supports for interconnecting various stakeholders in the knowledge sharing process (p 70)”.

Social support mechanisms such as story-telling (Kleiner and Roth 1997) communities of practice (Wenger 2000) are increasingly important given developments in the labour market. For example, the problem of aging in Europe could mean 20% of the European population will be over 65 years old by 2025 (Farrell 2005). Therefore, this means that knowledge of workers need to be harnessed effectively so that the impacts of knowledge loss amidst a rapidly retiring working population and a constricting labour market can be mitigated.
More crucially in the short-term, a recent study revealed that 90% of the UK workforce have permanent contracts; yet, these workers stay in the same organisation for merely an average of seven years and four months (Taylor 2002). With such trends, the harnessing of knowledge from the workforce becomes an even more pressing priority.

**Trust-based relationships**

The concept of trust means different things to different people. Inherent within the nature of the concept lies an element of ambiguity and complexity (Kidd et al. 2003) and several commentators have suggested that trust is multi-dimensional and multifaceted (Sako 1992a; Ganesan 1994; Fukuyama 1995; McAllister 1995). Farrell and Knight (Farrell and Knight 2003) define trust as ‘a set of expectations held by one party that another party (or parties) will behave in an appropriate manner with regard to a specific issue’ (p541).

Discussions on trust often occur at two levels. Game theoretic approaches focus on interactions between individuals. For example, Coleman (Coleman 1990) conceptualises trust in terms of relationships between ‘trustors’ and trustees whose decisions about whether to trust another are mediated by knowledge, probability of future gain from action of the ‘trustor’ towards the trustee, and calculations on loss associated with making an incorrect decision as to whether to trust another; by contrast, macro/meso level studies focus on institutions and rituals; for example, Putnam’s work on social capital defined as ‘[the] features of social organisation, such as trust, norms and networks, that can improve the efficiency of society by facilitating co-ordinated actions’ (Putnam, 1991 p167).

In the business context, both levels have much to commend them. The latter, however, has been used to analyse the most celebrated examples of trust-based business transactions – namely, the Italian industrial districts (Piore and Sable 1984). Where once the districts were seen as exemplars of ‘flexible specialisation’ and non-contractual co-operation between producers, suppliers and others, Farrell and Knight show how for institutional reasons these trust relationships are in decline as a mechanism for the delivery of productivity and global competitive advantage.

Previously, then, trust was built within the context of a community and sense of belonging and mutual interdependence. The community effectively regulated opportunism between cooperating firms and imposed sanctions on those who broke the rules, the ultimate penalty being expulsion from the community (Brusco 1992), pp182-83). Recent evidence of a breakdown of trust points to two key conclusions. First, that structural changes in the ownership of large ‘final’ or ‘governor’ firms has led to a formalisation of contract relationships stripped of the cultural context and significance (derived from local ownership, community and the need to ‘belong’). Moreover, this ‘depersonalisation’ leads to a much more contract-led mode of operation for firms in producer networks or value chains.

Essentially, the shift is from ‘networks of firms’ towards ‘networked firms’ – a subtle but important difference (Crouch, Le Gales et al. 2001) with access to customers being one of the critical resources for final or governor firms. The significance of trust for business organisations, despite the erosion of trust-based relationships in the Italian industrial districts arising out of institutional changes, remains an important component of innovation management in modern organisations. The importance of trust is emphasised in increasing technological collaboration (Jarillo 1988; Sako 1992b; Dodgson 1993), and quasi-legal structures that enable it such as joint ventures (Buckley and Casson 1988; Buckley and
Casson 1996). For example Saxenian’s (1991) study of Silicon Valley firms which involves ‘…relationships with suppliers as involving personal and moral commitments which transcend the expectations of simple business relationships’. (Saxenian, 1991 p428). Collaborative relationships are increasingly affected by cultural factors such as language, regional loyalties, educational experiences, ideology and even common leisure interests (Freeman 1990).

Complexity arising out of temporal and spatial separation demands considerable application and management. Lamming (Lamming 1993) worries himself with risks associated with choosing partners or collaborators for outsourced design capability. ‘The automotive industry’ he observes, ‘is a surprisingly close-knit community. If a supplier were to leak information of a truly strategic nature from one assembler to another, it would soon be known and the supplier’s credibility would be destroyed. Leakage in the other direction – from one supplier to another via an assembler – is more difficult to detect accurately but appears to be more commonplace.’ (p210) Interestingly, here, Lamming makes the link between trust and leakage.

**Alternative (future) view of trust**

Adler (2001) provided an alternative reading into Marxism and conceptually reviewed the concepts of market (i.e. price mechanism), hierarchy (i.e. authority) and trust (i.e. community) literature to illustrate the future of capitalism and the role of the knowledge economy. Adler (2001) proposed that the trend of high-trust institutional forms will proliferate in the knowledge economy (where growth in knowledge-intensity is a trait). However, [F]or trust to become the dominant mechanism for coordination within organisations, broadly participative governance and multi-stakeholder control would need to replace autocratic governance and owner control […] for trust to become the dominant mechanism for coordinating between organisations, comprehensive but democratic planning would need to replace market competition as the dominant form of resource allocation. (p230)

Adler (2001) qualified the notion of trust to state that it is not “blind” trust. That is, “its rationality is not of the purely calculative kind assumed by economics [rather] the values at work in modern trust are those of the scientific community: universalism, communism, disinterestedness, organised scepticism” (p227), so-called “reflective trust”. Adopting this notion of “reflective trust” could provide a plausible alternative to the model developed from the Italian industrial districts that, as discussed above, has seen the erosion of trust.

Although trust is considered to play an important part in encouraging people to share their knowledge, researching trust is also problematic. For social and political scientists, institutions often serve as the unit of analysis and the nature and type of power relations provide a proxy for trust. Invoking Farrell and Knight again, it is possible to investigate changes within and between organisations to explain levels of trust (often a case of bargaining between relatively powerful actors to achieve outcomes including co-ordination). However, the danger with this approach is to focus on compliance as a measure of trustworthiness to the exclusion of other less tangible indicators such as informal contacts, exchange of information and sanctions.

Finally, the need to avoid the trap of exclusively functionalist explanations for trust is important. Functionalist ontology has it that we explain phenomena by reference to their role in meeting objectives or realising outcomes. Thus trust is functional for mutually beneficial
co-operation and hence performance. It also reduces transaction costs associated with contracts, regulation and compliance (amongst other things). But the danger, according to Farrell and Knight (2003), is that we are overly reductive in that we simplify too much the causal mechanisms for enhanced performance and co-operation. Indeed, Farrell and Knight go on to suggest that the functionalist explanations are liable to be tautological. For example, that mutually beneficial co-operation is inherently good and by consequence it explains co-operation with reference to its mutually beneficial nature.

And finally we direct readers back to Ndofor and Levitas’ (2004) signalling quadrants, and the trust elements of their four quadrants; clearly trust is facilitated by low uncertainty, but highly endowed firms have much to lose from knowledge leakage which may be unintentional by inappropriate signalling.

**Links between organisational knowledge and productivity**

Lapre and van Wassenhove’s (Lapred and Van Wassenhove 2001) work on process improvement and their transfer to other units demonstrates the value of knowledge leakage and the practical problems associated with deploying it effectively. Indeed, they conclude that the next production frontier is the operation of factories as learning ‘laboratories’ defined as complex organizational ecosystems that integrate problem solving, internal knowledge, innovation and experimentation, and external information (Leonard-Barton 1992). The role of transferable knowledge is evident from the following example:…in the 1970s he had participated in an R&D project on the ability of tire cord to withstand corrosion. From this R&D project, he remembered that some copper-related variables determined in the brass coating step were relevant for the problem at hand in the WWD [wet wire drawing] department. The MLA [Model Line in Plant A] team tested the model with controlled experiments. As a result the MLA obtained a sharp improvement in productivity (p1316). What we didn’t understand when we started model lines in plants B and C is that a model line manager needs to have authority over production and projects, and a young engineer is not a good choice to run a model line. Young engineers lack experience in formal problem solving…(p1322).

Essentially, the success of the MLA owes much to embodied and accrued knowledge held by a key individual in combination with empowerment and good problem-solving skills. This is not always perfectly replicable. There are few other studies linking knowledge transfer and productivity explicitly, but we point readers to work a cross-sectoral study (based on questionnaire surveys) that indicate that firms in the UK do not actually measure their productivity because current measurements (via the economic perspective) are too complex, time consuming and labour intensive (Grugulis and Bevitt 2001).

Another study of note is Peter F Drucker’s thinking on knowledge worker productivity (Drucker 1999). (see also (CEC 2000) on the EU widely-known Lisbon strategy). There is equal – or even greater – opportunity in the developed countries to organize non-manufacturing production (i.e. production work in services) on the production principles now being developed in manufacturing. There is equally a tremendous amount of knowledge work – including work requiring highly advanced and thoroughly theoretical knowledge – that includes manual operations. Drucker points to six major factors that determine knowledge-worker productivity, five of which he believes differentiate knowledge workers from traditional production workers. These are (p83-84):
Knowledge-worker productivity demands that we ask the following questions of him or herself: What is the task? What should it be? What should you be expected to contribute? What hampers me in doing my task and should be eliminated? This contrasts with the production worker who merely asks, ‘how should the work be done?’

It demands that we impose the responsibility for their productivity on the individual knowledge workers themselves. Knowledge workers have to manage themselves. They have to have autonomy.

Continuing innovation has to be part of the work, that task and the responsibility of knowledge workers.

Knowledge work requires continuous learning on the part of the knowledge worker, but equally continuous teaching on the part of the knowledge worker.

Productivity of the knowledge worker is not – at least not primarily – a matter of the quantity of output. Quality is at least as important.

Finally, knowledge-worker productivity requires that the knowledge worker is both seen and treated as an ‘asset’ rather than a ‘cost’. It requires that knowledge workers want to work for the organisation in preference to all other opportunities.

What we can draw from these conditions is the realisation that (organisational) knowledge – if it exists at all – is a process rather than an artefact. The conduit is people, and whilst it can be deposited in some repository or captured in some software solution, more significant is its diffusion amongst people in its application to problem-solving across functions.

A more efficient productivity strategy is to share knowledge about up-to-date activity including process, change in product and services. This is mainly as it will have an impact on the employees in making them feel that they have an important part in making the organization successful.

**Literature synthesis**

From the selected literature reviewed, four emergent issues appear to surface. First, the concept of knowledge has tended to be used in the abstract. Arguably, this has led to greater confusion as the theoretical frameworks expand. To understand the relationship between knowledge leakage and productivity, we maintain, requires for greater specificity in terms of what knowledge is.

Second, studies abound in terms of linking knowledge and organisational performance, often with an emphasis on knowledge management and sharing. Whilst it is important to look at organisational performance, this does not necessarily inform us of the implications on organisational productivity – productivity should not be conflated with performance, but rather treated as a subset of performance, which forms the crux of the study.

Thirdly, the review highlights the need for a more holistic approach towards the understanding of knowledge and its impacts to productivity. As it stands, the literature is populated with studies from various perspectives in isolation. By developing taxonomy of knowledge leakage, this study attempts to synthesise more holistically our understanding of knowledge. Finally, the interplay between people issues with knowledge and productivity has been found to be weak in the literature. This is therefore a further gap that the current study seeks to address.
References


