Volume Title: New Developments in Productivity Analysis

Volume Author/Editor: Charles R. Hulten, Edwin R. Dean and Michael J. Harper, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-36062-8

Volume URL: http://www.nber.org/books/hult01-1

Publication Date: January 2001

Chapter Title: Service Sector Productivity Comparisons: Lessons for Measurement

Chapter Author: Martin Neil Baily, Eric Zitzewitz

Chapter URL: http://www.nber.org/chapters/c10131

Chapter pages in book: (p. 419 - 464)

10

Service Sector Productivity Comparisons: Lessons for Measurement

Martin Neil Baily and Eric Zitzewitz

10.1 Introduction

Output per hour in the nonfarm business sector grew at only 1.4 percent per year in 1973–95, with the weakness apparently concentrated in the service sector. Stagnant productivity in services seemed to belie the rapid adoption of information technology in this sector and the rapid pace of change in many service industries.

This puzzle, together with the parallel concern about measuring the cost of living, fueled efforts to improve price and hence productivity measurement. The measurement improvements that have been instituted at the Bureau of Economic Analysis (BEA) and the Bureau of Labor Statistics (BLS) resulted in an upward revision of 0.3 percentage in the rate of output per hour growth in the nonfarm business sector (from 1.1 to 1.4 percent). Moreover, starting in 1996 overall labor productivity growth accelerated sharply to around 3 percent per year, with the acceleration occurring in both service-producing industries and goods-producing industries (see the disussion in Council of Economic Advisers 2001).

There remain questions, however, about how well service sector output is being measured, and these questions apply to services sold both for final use and for intermediate use. Measurement errors in the former affect estimates of both industry level and overall output and productivity, whereas errors in the latter affect only our understanding of productivity by industry. Despite the overall improvement in service sector

Martin Neil Baily is a Senior Fellow at the Institute for International Economics. Eric Zitzewitz is a graduate student at the Massachusetts Institute of Technology. They were, respectively, principal and economics analyst at the McKinsey Global Institute.

The authors would like to thank Robert J. Gordon and other participants in the conference.

productivity growth, data anomalies remain, including sectors where reported productivity is declining.[1]

Starting with an international comparison of service sector productivity published in 1992, there have been a series of projects initiated by the McKinsey Global Institute and carried out in collaboration with McKinsey & Company offices worldwide and a number of academic economists.[2] Although the emphasis of these projects has been to measure and explain productivity by industry by country for a range of service sector and goods-producing industries, another important concern has been with employment creation and the extent to which productivity increases will cause or alleviate unemployment.

The purpose of this paper is to ask whether the results of these international comparisons cast any light on the problem of service sector output measurement. To do this, we present case studies of five service industries: retail banking, public transport, telecom, retailing, and airlines. We describe how we have measured output and what we believe are the main explanations for the resulting cross-country productivity differences. To make one point clear at the outset, these studies have not generally developed new output measures. The basic output measures in the five cases are the BLS index of retail banking output; vehicle kilometers in public transport; access lines and call minutes in telecom; value added in retailing; and revenue passenger kilometers (RPK) for airlines. In selecting output measures, the goal has been to find metrics that are intuitive and understandable to noneconomists, and business consultants in particular, and that are at least acceptable to economists. Where possible, we have had a strong preference for physical output measures in goods-producing sectors, such as liters of beer. We also have used similar tangible output measures for services, such as call minutes of telecom services used.

Where our efforts may help improve the measurement process is by providing an understanding of how these industries have evolved (because the cross section allows us to see industries at different stages of evolution); how innovations have been used to generate customer value; and what has caused the productivity differences.

In industries where the share of purchased inputs in output is large and may differ significantly across countries, we have used value added as the output measure, and this applies to the retail case we describe later. Value

1. There is concern that efforts to improve measurement may be biased because most researchers are looking for reasons that output and productivity are understated, whereas in practice it is possible that quality has deteriorated in some service industries. We are aware of this problem, but in general we find that successful service sector innovations must drive out existing providers, so the presumption is that industry evolution is generating consumer surplus as people choose the new combination of price and service quality.

2. See the reference list for the specific citations. The number of people working on these projects is too numerous to mention each individually. Bill Lewis, the director of the McKinsey Global Institute, has supervised all of the projects.

added is not the favored output measure among productivity economists, who prefer to use a production function with gross value of production as the output and materials as one of the factor inputs. The McKinsey analyses do use the production function as the central conceptual framework for explaining productivity differences at the production process level, but we do not estimate formal production functions. The concept of value added is intuitive to business consultants whose own work is often based on the value chain concept. Where possible we try to carry out a double deflation estimate of value added, using purchasing power parity (PPP) for inputs as well as outputs. That is often hard to do, and the estimates are subject to error.

10.1.1 Quality

Because our analyses have largely been of cross-sectional comparisons at a point in time, we have been able to avoid the problem of changing quality over time. Where, based on industry knowledge, we find significant quality differences across countries in a given year, we try to control for those differences. For example, in autos, we had access to conjoint studies carried out in the United States and Europe that estimated the quality premiums of Japanese and European nameplates over U.S. nameplates.[3] In services, a recent example of quality adjustment for telecom service in Brazil had to account for the difficulties of noisy connections as well as the inability to place a call. In this case we estimated the investment in equipment that Brazil would have to make to bring its service quality up to the U.S. level, and we effectively reduced capital productivity and hence TFP to reflect this adjustment. A purer alternative would have been to estimate the price adjustment that would have made consumers indifferent to the different service levels (the conjoint approach). We did not have the data to do this. Moreover, the answers would likely have been very different for the United States and Brazil.

The quality adjustment is not always done in an unbiased way—for example, suppose we find that Country A has productivity that is twice as high as Country B using a raw output number, and that service quality is somewhat higher in Country A. Lacking a quantitative adjustment that can be made with reasonable resource cost, we may decide to ignore the quality difference on the grounds that it will not change the overall analysis of the causes of the productivity differences much.

10.1.2 Interaction between Measurement and Causality

A basic tenet of scientific research is that data collection should be independent of the hypothesis being tested by the data. A double-blind clinical

3. This approach is similar to the hedonic approach pioneered by Griliches.

trial of a new drug is a standard example. Economists try to follow the same rules by using standard government data or by collecting data in an objective or arms-length manner. Though sensitive to the demands of rigorous hypothesis testing, the McKinsey studies have not operated in the same way. We are not generally testing some aspect of economic theory. Instead, we are trying to provide our best estimate of both the *magnitudes* of the productivity differences among industries in different countries and the *explanations* for those differences at the production process level. When a McKinsey team working on one of the case studies reports to us that there is a productivity gap of, say, 40 percent between the U.S. and another country, we do not accept that figure until the team has made plant visits and has talked to experts in the industry and found supporting data to explain at the operational level why the productivity gap occurs. What is different about the check-clearing process, for example, that can account for a large labor productivity difference across countries.[4] Sometimes, in that process, we discover that substantial adjustments must be made to the productivity numbers. For example, we may find that the scopes of the industries are different in the countries, or that some workers employed by one industry actually work in another industry.

The drawback with this approach is that our analysis of causality at the production process level is not true hypothesis testing. The advantage is that we believe it provides much greater reliability. Subsequently, we do attempt to test hypotheses at a different level. Having determined the productivity gaps and the reasons at the production process (or production function) level, we then ask the higher-level question of why companies operate differently in the different countries.

The paper goes through each of the case studies in turn, attempting to draw out the lessons for measurement. We conclude with a review of the findings.

10.2 Retail Banking

Retail banking has undergone rapid changes in its competitive and regulatory environment since the early 1980s, and the fact that these changes have varied in speed and intensity across countries makes it an interesting industry in which to study their effects. In the United States, the deregulation of interest rates and money market checking accounts in the early 1980s increased competitive pressure on retail banks. The increased pressure together with an increasing application of IT coincided with an acceleration in labor productivity growth (as measured by the BLS, fig. 10.1).

4. Because our focus is on productivity comparisons, our method does not generally identify ways in which best practice could improve. Our use of the term *best practice* does not imply that we think such improvement is impossible.

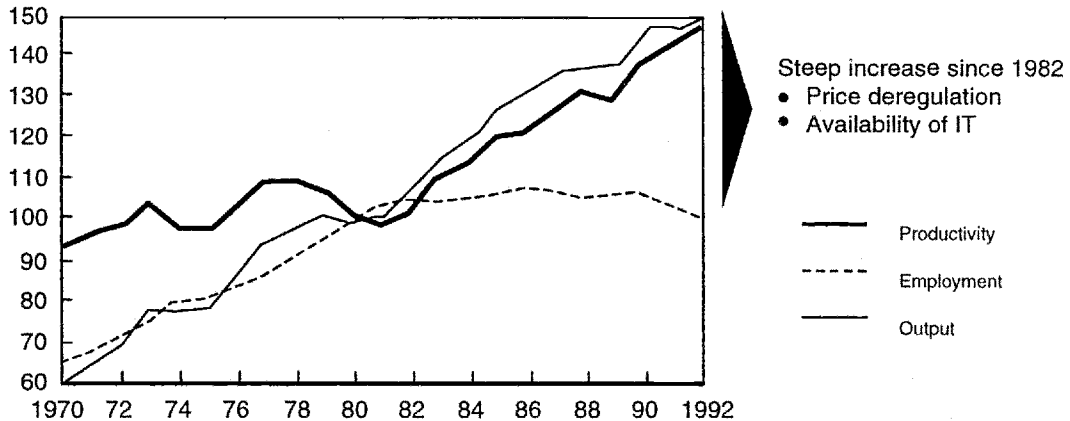

Fig. 10.1 Labor productivity, output, and employment: U.S. commercial banks (Index: 1980 = 100)

Note: Includes both traditional products and mortgages, FTEs

Source: Output and productivity data provided by the Bureau of Labor Statistics

In the late 1980s and early 1990s, competition and IT use were further increased by the emergence of specialized processing and mortgage/consumer loan companies, the liberalization of interstate banking, and the emergence of a market for corporate control.

Outside the United States, the retail banking industry is typically more concentrated, and the increase in competitive intensity in retail banking has been less dramatic. In most countries, lower competitive intensity has been accompanied by lower productivity, but the Netherlands has achieved the highest productivity of the countries we have studied with highly concentrated sectors. High concentration may have contributed to high productivity by aiding in the development of an electronic payments system and by reducing the pressure for expanding the branch network.

10.2.1 Productivity Measurement

The recent changes and cross-country differences in the banking industry complicate productivity measurement. The traditional banking products included in the BLS output index (numbers of transactions and deposit and loan accounts) are easier to capture in physical measures than are the newer products into which retail banks are expanding. Recent changes in technology have also increased the cross-country and time series differences in the convenience with which banking products are offered.

Measuring Output

We measure the overall output of traditional banking products using a methodology similar to that of the BLS. We use our banking practice knowledge to divide the employment[5] involved in producing traditional retail banking products into three functions: processing payment transactions, maintaining deposit accounts, and lending. We then construct functional productivity indexes and an aggregate productivity, which is the average of the functional productivity measures weighted by employment.

This yields the same results as weighting outputs by the labor required to produce them.[6] This is the same procedure we use in the other service industries with multiple outputs: airlines (RPK and passengers moved) and telecom (access lines and call minutes). Using labor requirements as weights is necessary because these outputs are usually not separately

5. Ideally we would use total factor input in measuring productivity, but measures of the physical capital involved in producing retail banking services (as distinct from the other assets of a bank) are very difficult to construct and make consistent across countries.

6. To see this, consider the following expression for countries A and B and products C and D: $(Q_C^A/L_C^A)/(Q_C^B/L_C^B)*[L_C^A/(L_C^A + L_D^A)] + (Q_D^A/L_D^A)/(Q_D^B/L_D^B)*[L_D^A/(L_C^A + L_D^A)]$, which is the average of the two product-level productivity ratios, weighted by the labor share of the products in country A. This is equivalent to $[Q_C^A*(L_C^B/Q_C^B) + Q_D^A*(L_D^B/Q_D^B)]/(L_C^A + L_D^A)$, where the products of country A are weighted using the unit labor requirements in country B.

priced in a way that reflects their true cost. Just as airline passengers do not pay separate per flight and mileage charges, banking customers pay for their deposit accounts in the form of an interest margin and usually receive their transactions for free.

Defining Output: Volume versus Value, Stock versus Flow

In following the BLS output measurement methodology, we have made two important decisions about measuring output in banking that are worth reviewing. We measure the number, not the value, of transactions and accounts under the assumption that both the customer value and bank input requirements are more constant per transaction/account than per dollar value. This is obviously not completely true for any product, especially for lending, where large loans typically are reviewed more carefully and provide a greater service to the borrower than do smaller loans. We compensate partially for this problem by weighting loan and deposit accounts by the cost factors estimated in the functional cost analysis, thus capturing the fact that maintaining an average mortgage requires more work than does maintaining an average credit card account, but not the fact that high-value mortgages can require more work than smaller ones.

Another decision we made is to measure the stock of deposit and loan accounts outstanding, rather than the flow of new accounts. Because the turnover of accounts is more rapid in the United States than in other countries and has become more rapid over time, this decision may understate both the level and growth in U.S. productivity. Failing to capture the higher account turnover is more justified in some cases than in others. Some mortgage refinancing is probably not positive net economic output but can be viewed as a particularly inefficient way of adjusting a fixed interest rate to market conditions. At the same time, some of the shorter average life of mortgages in the U.S. is caused by higher mobility; the accommodation of this mobility by the mortgage industry through a higher loan turnover ratio probably does reflect additional output. Thus far, we have only addressed this issue by adjusting downward the numbers of deposit accounts in countries that have a high share of unused dormant accounts.[7]

Unresolved Issues: New Products and Convenience

In response to the increasing offering of nontraditional banking products by retail banks, we attempted in a couple of projects to include physical measures of these products, such as the number of equity transactions or mutual fund accounts. We abandoned these measures mainly because

7. This adjustment was originally motivated by the fact that in Sweden customers average eight deposit accounts per capita (compared with 1.5 in the United States), mainly because Swedish banks do not attempt to discourage low/zero-balance dormant accounts.

they produced implausibly low productivity levels for the U.S. securities industry. About half a million people work in the U.S. securities industry, most of whom are involved in functions other than the processing of transactions and the maintenance of mutual funds. These functions are present to a much greater degree in core financial markets like the United States and United Kingdom than in peripheral markets, but we have not found a methodology for adjusting either the output or input figures to reflect this difference. Other nontraditional banking products (such as insurance) present similar problems.

Capturing differences in the convenience with which banking services are provided is another unresolved issue in these productivity comparisons. The increasing penetration of ATMs, PC banking, and other delivery channels has obviously increased convenience greatly over time. Fortunately for our cross-country productivity measures, these changes have occurred at a roughly similar pace throughout the advanced countries, and our banking practice judged the level of convenience offered in transaction and deposit products to be roughly similar across countries.

Important differences in convenience do exist in lending, however. Information technology has allowed lenders to centralize lending decisions, improving the quality and timeliness of the decisions. Mortgage and other loan decisions are now made in minutes rather than days, and better information has allowed lenders to extend credit to more risky borrowers than in the past. This technology was initially developed in the United States by specialized lenders such as Countrywide Credit, whose growth helped encourage its adoption by the rest of the U.S. industry. Outside of the United Kingdom, this technology has penetrated other countries more slowly, and thus the convenience associated with rapid lending decisions (i.e., greater availability of credit) is greater in the United States. We have not yet developed a good methodology for capturing this extra convenience.

10.2.2 Productivity Results

Overall we find the Netherlands to have the highest productivity of the countries studied, with the United States and France tied for second (table 10.1). The Netherlands leads the United States in all three functions, but especially in processing payments. The Dutch advantage comes from the combination of a more efficient electronic payments system and an efficient branch network (table 10.2).

After adjusting for these two factors, the United States actually uses its labor more efficiently, in part due to more centralization of processing and use of part-time labor in the branches.

Electronic Payments

The development of the electronic payments system in the Netherlands was led by the Postbank. When the Postbank introduced its electronic

Table 10.1 Labor Productivity in Retail Banking, 1995 (Index: U.S. = 100)

Country	Productivity	Country	Productivity
Netherlands	148	Australia (1994)	60
United States	100	Brazil	40
France	100	Colombia (1992)	30
Germany	85	Mexico (1992)	28
Sweden (1993)	80	Venezuela (1992)	25
Korea	71	Argentina	19
United Kingdom (1989)	64		

payments system, the other banks fought back by launching a joint alternative electronic payment system of their own. Their collaborative effort facilitated the development of standardized payment specifications that are essential for efficient electronic payment systems. In the Dutch case, the banks' joint efforts led to the establishment of a unique IT format adjusted for small payment transactions, a nationwide account numbering system, and convenient payment vehicles such as electronic debit cards. The end result has been widespread acceptance of electronic payments, with high levels of efficiency and customer convenience in payments. Other countries such as Germany and France also have Dutch-style electronic payments systems, but these countries have lost this advantage through dense branch networks (Germany) and inefficient labor utilization.

The United States has had difficulty setting up shared electronic payments systems for two key reasons. First, the highly fragmented nature of the U.S. retail banking industry, as well as state laws that have limited the interstate activities of banks, have impeded the development of common standards for IT and account numbering. Second, customers have been reluctant to switch to electronic payments; Americans continue to rely on checks for 70 percent of all transactions, perceiving them as safer and better suited to record keeping than other payment vehicles. However, in the countries that have recently started shifting over to electronic payments, such as the United Kingdom and New Zealand, customers have been given a small incentive to shift and have responded to that incentive. We believe there is a substantial unexploited opportunity for the United States to raise banking productivity and, once the transition has been made, to increase customer convenience.

Branching

The consolidation of the branch network in the Netherlands is helped by that country's high population density and concentration. The density of the Netherlands allows the banking industry to have 2.5 and 6 times more branches per square kilometer of residential area than France or the United States, respectively, despite having 45 percent and 30 percent fewer branches per capita. The Dutch banks have not squandered this natural

Table 10.2 **Drivers of Productivity in Payment Services (Index: U.S. = 100)**

	Netherlands	France	Germany	United States	Cause of High Performance
Share of electronic payments	175	154	182	100	Joint standard setting
Branches per capita	119	93	79	100	High population density
Average labor utilization	74	56	54	100	Labor Organization
Overall payments productivity	154	80	78	100	

Source: McKinsey Global Institute and the Max Geldens Foundations for Societal Renewal 1997.

advantage (as the German and Korean banks have) by overexpanding their branch network, in part because the high levels of concentration reduces the pressure to compete through branch proliferation. The absence of regulations or understandings on pricing has also reduced the incentive to focus on competing through branching.

Labor Organization

The U.S. industry has been most successful relative to other countries in organizing its labor efficiently. The U.S. industry makes much greater use of part-time labor to staff efficiently for daily and monthly activity peaks in the branches. Using part-time labor is more difficult outside the United States because of collective bargaining agreements or semivoluntary adherence to traditional labor practices. Greater penetration of IT in the United States also allows the lending decision making, transaction processing, and account maintenance functions to be centralized.

10.2.3 Lessons from the Banking Case

Two main implications for measuring productivity growth and levels can be drawn from the banking case. The first is that there have been significant increases in the efficiency of banks associated with the increased penetration of IT. Our industry practice views the pace of change as being at least as rapid as in manufacturing industries. At present the BEA effectively assumes away labor productivity growth in all financial services (changes in real output over time are set equal to changes in full-time equivalent persons engaged in the sector). They could certainly improve on that for retail banking by making use of the BLS banking productivity estimates (which show productivity growth of 4.5 percent per year since deregulation in 1982). The second implication is that even the BLS estimates may be missing changes in the true output of the sector. Some of the increase in the turnover of accounts in the United States probably represents increased true economic output; both the BLS and our measure of output miss this increase. More importantly, the increased convenience and availability of lending is not captured in these estimates.

10.3 Telecom

Our telecom case studies have all focused on the traditional, fixed-line residential and commercial services, which still account for about 80 percent of telecom revenues. Data for mobile phones are included in the analysis, raising the coverage of the case study to roughly 90 percent of revenue, but the overall results are still dominated by the fixed line business.

The telecom industry essentially provides two outputs. The first is access lines, which provide access to the telephone network and the option of making and receiving calls, and the second is the actual call minutes them-

selves. These two outputs are priced and consumed at very different relative levels across countries, and this complicates attempts to measure productivity. For example, in Brazil the supply of access lines was restricted by high import tariffs on capital goods and the government's diversion of telecom earnings to finance budget deficits. As a result, access lines per capita are very low given income levels (the black market rate for a line has reached $5,000), but use of these lines is very high. In other countries, such as Japan and Germany, access lines are easily available and monthly fees are low, but calls are expensive and use is low. Aggregating the two telecom outputs into a common productivity measure is one of the major challenges of the case.

10.3.1 Productivity Measurement

Telecom is one of the three case industries with multiple physical outputs where the outputs are not separately priced in a way that reflects their true cost or value to the customer. As in the other industries, we attempted to aggregate the outputs using weights based on the inputs required to produce them. Unlike banking, data for telecom is available for both labor and capital input. Labor input in telecom is used primarily to maintain and install access lines; only about 15 percent of the labor is used in activities where labor requirements are determined by call traffic. Capital, however, is mainly the switching equipment and lines that are used to provide call minutes.

As in banking, we constructed our labor, capital, and TFP measures by weighting outputs according to their input requirements. Labor productivity was measured as a weighted average of access lines maintained and call minutes per hour worked, and capital productivity was measured as call minutes per unit of network capital services.[8] Labor and capital productivity were then averaged to form TFP, using weights based on the labor and capital shares in value added. The advantage of this procedure is that it produces both an acceptable overall measure of productivity as well as component labor and capital productivity measures that are meaningful to our industry audience.

Capital services are constructed using a perpetual inventory method with sudden death depreciation. Capital is divided into four categories (switching equipment, cable/wire, land and buildings, and general purpose) and assigned service lives based on asset lives given by the Federal Communications Commission (FCC) to the Regional Bell Operating Companies. Capital expenditures are converted into U.S. dollars using PPPs

8. The capital involved in maintaining the network was impossible to separate from the network capital but is a relatively small share of overall capital. Because the network maintenance–related capital was included as providing call minutes, we may have slightly underweighted access lines relative to call minutes in constructing our output measure.

for structures, equipment, and telecom equipment. That capital goods tariffs caused access lines to cost $2,700 in Brazil instead of $2,300 does not lower Brazilian capital productivity as measured. But Deutsche Telekom specifications, which raised capital requirements 20 percent by calling for main wires to be "tank proof" (i.e., able to survive being run over by a tank) even when they are placed underground and requiring the use of underground as opposed to aerial cables, do reduce German capital productivity as measured. Capital productivity is also adjusted for service quality. In Brazil, where service quality is lower due to the use of mechanical and analog instead of digital switches and to lower switching capacity per access line, we estimate the capital that would be required to upgrade Brazilian service to U.S. standards (about 5 percent more capital per line) and lower Brazilian capital productivity by this amount.

The main issue with using input requirements as a method for valuing output is that they may not reflect the value of the output to the consumer. The high black market price for access lines in Brazil suggests that adding one million lines in Brazil would be more valuable than adding one million lines in the United States. Likewise, a German consumer who is paying four cents per minute for a local call is probably getting more value from his or her average call minute than is the American with free local calling whose teenagers gossip for hours to their friends.

To give one check of whether calls in the United States are of lower value than calls in other countries, we looked at data on average call length. We found that the average call length in the United States is about the same as in Europe. Gossiping teens do not explain the usage differences. Moreover, one point of view is that there are externalities associated with telecom usage, where it drives productivity improvements elsewhere in the economy. Another check of the results was carried out by the McKinsey team in the United Kingdom, where a study was being carried out as this paper was written. The team examined the share of telecom revenues accounted for by payments for the access line and payments that are determined by call-minutes. Data were collected for the United States, United Kingdom, Germany, France and Sweden. As expected, the United States has the largest share of revenues from access line charges (35 percent), but the other countries' shares are not all that different. The lowest is France, at 20 percent, where local calls are very expensive. The others are Sweden 34 percent, United Kingdom 27 percent, and Germany 26 percent. Using the average revenue share of 28 percent to weight lines and minutes gave productivity results that were pretty much the same as the ones reported here.

The countries in which we have studied telecom productivity can be roughly divided into three groups (table 10.3). The United States is the only country studied that had both high labor and capital productivity, which mainly reflect high access lines per employee and high network utili-

Table 10.3 **Productivity in Telecom (Index to U.S. = 100, 1995)**

	United States	Germany	Korea	Japan	Argentina	Brazil
Total factor productivity	100	42	66	49	56	64
Labor productivity	100	51	82	74	70	45
Access lines per FTE	100	83	88	80	75	46
Call minutes per FTE	100	39	57	33	36	39
Capital productivity	100	38	58	40	51	75
Access lines per dollar of capital services	100	61	65	42	48	86
Call minutes per line	100	48	112	106	90	109
Call minutes per capita	100	37	40	30	14	12
Access lines per capita	100	79	62	71	28	14

Source: McKinsey Global Institute (1998a, b).

zation, respectively. The United States had thus efficiently (relative to other countries, at least) built and maintained a large network and achieved high usage of that network. Most of the other countries studied have achieved fairly high labor productivity but low network utilization. Brazil, on the other hand, had low labor productivity, which resulted in an insufficient supply of access lines and thus high network usage. Before privatization Argentina actually suffered from both low labor and low capital productivity, although labor productivity has recently improved.

Labor Productivity

Cross-country differences in labor productivity are due to differences in the adoption of advanced technology such as digital switching and to differences in the pressure to rationalize employment in order to take advantage of the new technology. In Brazil less than 50 percent of access lines are digitally switched, compared with over 80 percent in the Organization for Economic Cooperation and Development (OECD) countries. Digitally switched lines require one-fifth as much maintenance technician time as analog lines to maintain, and Brazilian regional telecoms with higher digitalization had lower cash (non capital, i.e., mainly labor) operating costs. In addition, less IT is used in Brazil in directory assistance and customer service, which also leads to lower labor productivity.

New technology is important in enabling higher labor productivity, but in countries that have adopted digital switching and other modern technologies, labor requirements per line have declined rapidly. In the United States, for example, access lines per employee have grown at 7 percent per year since 1980. In most cases this potential productivity growth is more rapid than the growth in demand for access lines, and telecoms therefore have to reduce employment in order to capture the full benefit of the new technology. In general, we have found that private telecoms facing at least limited competition have been more likely than state-owned telecoms to make these reductions. For example, the one private regional telecom in Brazil has labor productivity that is 30 percent above average despite an only average level of digitalization.

Capital Productivity

The most important factors explaining differences in network usage and thus capital productivity are pricing and the availability and marketing of demand-enhancing call completion services. The marginal pricing of local, long-distance, and international calls is much lower in the United States than in Germany, Japan, and Korea, and demand for calls per capita is much higher. Most U.S. customers pay a higher monthly fee and receive free local calls, whereas most other countries do not offer free local calls as an option. In addition, services that increase call origination and termination are more available and more aggressively marketed in the United

States. Business call centers, voice mail, call waiting, and call forwarding have much higher penetration in the United States than in other countries, and the U.S. telecoms do not tax answering machines as Deutsche Telekom does. Telecoms in the United States have historically expended much more effort and money marketing telecom services, whereas in other countries there has been less focus on stimulating demand. In Germany advertising has even carried the opposite message: Several decades ago Deutsche Telekom ran advertisements advising consumers to "Fasse Dich kurz"—keep it brief.

10.3.2 Lessons from the Telecom Case

The differences in the pricing and availability of telecom outputs across countries provides evidence of how large the consumer surplus associated with telecom services can be. Even when the rights to an access line were trading at $5,000 in Brazil, there were seven lines per hundred population, and almost all consumers could have sold their line on the black market since the rights to a phone line were transferable. Thus roughly 7 percent of the population valued a phone line at over $5,000 in a country with 22 percent of U.S. GDP per capita. This suggests that the consumer surplus generated by the access lines that are available in the United States is significantly larger than the cost of building and maintaining the lines. Likewise, demand for local calls appears fairly inelastic above one to two cents per minute; even though the price of local calls is 5 cents per minute, U.K. consumers still demand over 1,300 minutes per capita. This suggests that the consumer surplus from call minutes is also large.

We have not even begun to attempt to measure this consumer surplus, although others have (e.g., Hausman 1997). The evidence in our case studies on pricing and demand across countries, however, suggests that understanding consumer surplus is important to understanding both telecom's final output and its contribution to the output of other sectors.

10.4 Retailing

In a series of studies, McKinsey Global Institute has examined cross-country comparisons of productivity and employment for general merchandise retailing, food retailing, and a combination of both. In this paper we will draw on the work that was done in the Netherlands, which covered all retailing. It is not ideal to combine food and nonfood, although given the presence of hypermarkets in Europe, it can be difficult to separate the two.

The importance of the retail sector to the national economy is often underestimated. Yet, it accounts for 5 to 6 percent of GDP and 7 to 11 percent of employment in the United States and Europe. It is particularly important in creating jobs for groups with high unemployment levels, employing relatively large numbers of women, young people, and people with little education. The sector is also a major provider of part-time work.

Because of this, our studies of retailing have often focused on employment. The existence of very high minimum wages in Europe (effectively about $8 an hour in France, given payroll taxes), together with zoning restrictions that inhibit the development of high service retailing formats—notably shopping malls with department store anchors and specialty chains—has adversely affected the creation of low-skill jobs.

10.4.1 Productivity Measurement

Format Evolution

Retail has undergone a major structural shift in the late twentieth century, as traditional stores have been partially replaced, first by department stores and then by specialized chains, mass merchandisers, and mail order. In order to understand the effects of this structural shift on productivity and employment in the sector, we have segmented the industry into six different store formats: (a) mass merchandisers, such as Safeway or Wal-Mart; (b) out-of-town specialized chains, such as IKEA and Home Depot; (c) in-town specialized chains, such as The Gap, Benetton, and The Body Shop; (d) department stores, such as Saks, Bloomingdales, and Nordstrom; (e) mail-order companies, such as Lands' End; and (f) traditional stores, such as bakeries and small hardware stores. E-commerce dealing was not significant at the time of the study.

These different store formats offer different value propositions to their customers. Mass merchandisers and, to a lesser extent, out-of-town specialty chains and mail order compete mainly by offering low prices (table 10.4).

In-town specialty chains and department stores compete more by offering high service levels, for example in the form of a targeted selection of merchandise. The traditional stores that remain survive by offering high service, usually in the form of convenient locations. In general, mass merchandisers and specialty chains are more productive than the other store types, and in both the United States and Europe the industry is shifting toward these formats.

Measuring Retail Service

The coexistence of in-town specialty chains with mass merchandisers despite value added to sales ratios that are over twice as high can only be explained if these specialty chains are providing extra service that customers value.[9] Given that the level of service provided by retailers can differ

9. Part of the difference in the gross margin-to-sales ratio can be explained by the fact that mass merchandisers specialize in fast-moving, lower margin goods, but there are also significant cross-format differences in prices for the same categories or even the same items. One can usually find items more cheaply at mass merchandisers like Wal-Mart, but many prefer a shorter drive and the greater availability of informed sales personnel in specialty chains.

Table 10.4 **Productivity, Throughput and Service Level of Retail Formats (Index: Mass merchandisers = 100)**

	Mass Merchandisers	Out-of-Town Specialty Chains	In-Town Specialty Chains	Department Stores	Mail-Order	Traditional Stores
Gross margin per hours worked	100	100	80	90	80	60
Sales per hour worked	100	70	35	55	70	25
Gross margin per sales	100	140	230	160	110	240

Note: Index is based on U.S. data. Traditional stores attain high margins partly by locating in areas underserved by more advanced formats.
Source: McKinsey Global Institute and the Max Geldens Foundations for Societal Renewal (1997).

so much, we rejected a simple throughput measure of retailing output. In principle, what we would like to measure is the value added[10] of retailers at some set of international prices. But this is complicated. Retail prices are available for food and general merchandise goods from the OECD and International Comparisons Project (ICP) PPP projects, but these retail prices are not usually collected in comparable formats (in part because comparable formats do not always exist), and wholesale prices are harder to obtain. As a result, double-deflated PPPs that truly reflect the cost (in terms of the gross margin a consumer must pay) of retailing service are essentially impossible to construct.

Because the ideal measure was not feasible, we considered two alternatives.[11] The first was to use the PPP for consumption as a measure of the "opportunity cost" of consuming retail service vis-à-vis other consumer goods and services. The rationale for using the consumption goods PPP is that the general price level of all consumption goods should give one at least an unbiased estimate of the price level of a specific consumption good, retail service. The problem with this method was revealed when we studied countries with very high relative retail wages (at PPP), such as France and Sweden. High retail wages increase our measured labor productivity for two reasons. The first is that high wages encourage retailers to forego providing marginal services, such as grocery bagging.[12] By eliminating jobs with low marginal labor productivity, average labor productivity should rise; this effect we would like to capture in our measure. Unfortunately, high relative retail wages also force up the relative price of retail service, which causes our measure to overstate retail productivity. Because we cannot construct a retailing service PPP, we cannot correct for this effect.

The retailing study made in the Netherlands attempted to overcome this problem by using, essentially, a same-format PPP. It is based on the assumption that the absolute productivity level of out-of-town specialized chains is similar in the Netherlands and the United States, because the retail concept applied in this format is very similar in the two countries. This assumption implies that the service and efficiency levels (value added per good and the number of goods sold per hour) are comparable, as is productivity. We used the absolute productivity level of out-of-town spe-

10. For retail, gross margin and value added are essentially equal. In the exposition of the methodology, we use the terms interchangeably in order to get concepts across more easily. In calculating our results, we mainly used data on gross margin because it is more widely available.

11. Actually, the Dutch team also used a third method that we are not reporting that adjusts for format mix only. See McKinsey Global Institute and the Max Geldens Foundations for Societal Renewal (1997).

12. One might have expected retailers to respond to high wages by substituting capital for labor, but because structures and rent (the main component of capital in retailing) are also more expensive in Europe, the main difference is the absence of marginal services in Europe.

cialized chains as a bridge and combined it with relative format productivities within each country (captured by relative values added per hour for the different formats within each country) to calculate absolute format productivities. We calculated retail sector productivity in each country by weighting the absolute format productivity with the share in retail employment of each format. This method actually calculates relative productivities rather than retail outputs, but there is an implied relative output figure computed by multiplying retail sector productivity by the number of hours worked. The implicit assumption is that when a U.S. Toys "R" Us employs more workers than a European Toys "R" Us to generate the same quantity of sales, it is producing 30 percent more retailing service. Toys "R" Us does provide more service in the United States; its opening hours are longer; and check out lines shorter. However, it may be an overestimate to say that extra employees translate into extra service one for one.

Unfortunately, the data were only available to apply this method to the U.S. comparison with the Netherlands. This approach suggests a higher relative productivity in the United States than does the regular PPP comparison. It may go too far and overstate the value of customer service in the United States, but it does provide a new estimate of the quantitative impact of the service component.

10.4.2 Productivity Results

The productivity results using the OECD PPPs indicate that retailing productivity is pretty much the same among the European countries and the United States. Using OECD PPPs, the productivity at the U.K., U.S., France, and German industries is 108, 105, 101, and 101, respectively, with the Netherlands indexed to 100. This similarity reflects two offsetting forces. The United States has fewer low productivity traditional or mom-and-pop stores (fig. 10.2). Its low relative wages in retail allow for more marginal services to be provided and may lower the relative price of retail service, causing our value added at consumption PPP to underestimate U.S. productivity.

The result of the United States–Netherlands comparison using the format bridge approach raises U.S. relative productivity by about 15 percent. This could be an overstatement of the difference. As we noted earlier, comparable formats tend to have higher staffing levels in the United States, and the bridge assumption implicitly says that the additional staff have the same average productivity as the intramarginal employees. That likely overstates their productivity, and the true U.S. productivity figure is probably between these two estimates.

Although productivity levels are fairly similar in the United States and Europe, U.S. retailing output and employment are 50 to 80 percent higher. Higher output is only partly due to 20 to 40 percent higher U.S. consumption of retailed goods. The U.S. format mix is shifted toward more service-

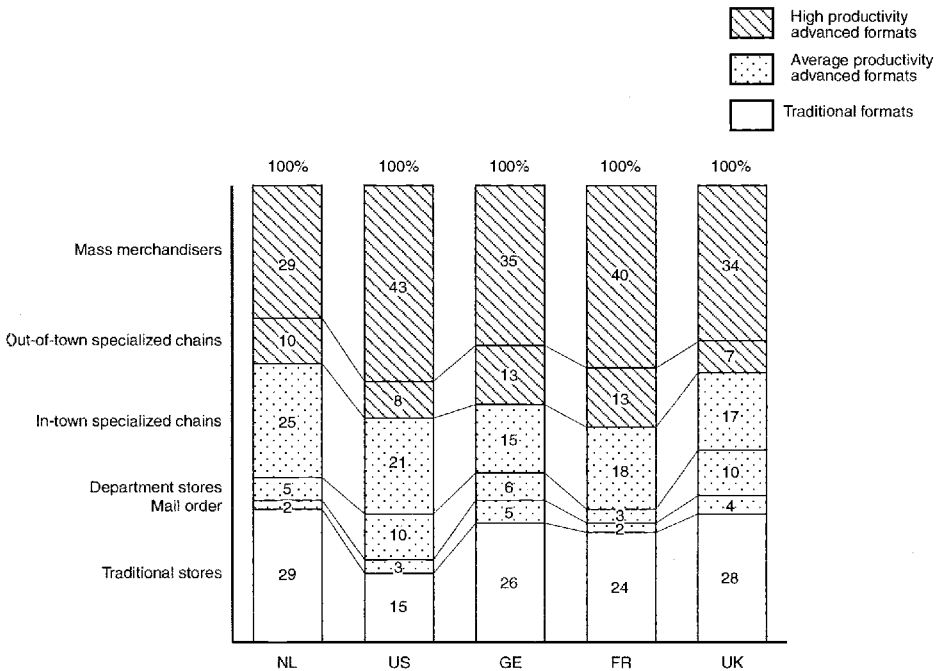

Fig. 10.2 International comparison of market shares of retail formats, 1994 (percent of total sales)

Source: McKinsey Global Institute and the Max Geldens Foundation for Societal Renewal 1997

intensive formats such as in-town specialty chains, and U.S. stores provide more service in a given format—in the form of longer opening hours, more sales assistance, shorter checkout lines, and so on. The less output and employment-intensive European format mix is partly due to zoning laws that prevent the development of shopping centers where in-town specialty chains can be successful. The lower employment and service levels within given formats is mainly due to high minimum wages and social costs, which cut off the lower edge of the employment distribution. This results in fewer open checkouts, no queue busters, shorter opening hours, less cleaning of the stores, fewer customer representatives, and no bag packers.

10.4.3 Lessons from the Retailing Case

The McKinsey retail practice told us back when the original service sector study was carried out some years ago that this industry in the U.S. is highly competitive and innovative. Aggregate data that suggested weak productivity growth in retailing seemed implausible to them and suggested that the data may be missing the shift in U.S. retailing employment toward specialty and higher service retailing formats where throughput per employee is low. The most obvious sign of this is the rapid growth of shopping malls. Ever since, we have been trying to capture the productivity differences that occur in cross-country comparisons because of the very different evolutions of retailing in the United States and other countries.

We believe that the format bridge approach is an appealing way to approach the problem. The PPPs are based on efforts to compare like goods to like goods. We have tried to compare like retailing service to like retailing service. As long as these bridges are competing in a major way in markets in the different countries, they can provide a PPP equivalent.

The lesson for the measurement of output and productivity over time in the United States is not a new one. Retailing is a service industry, not just a conveyor belt to bring the goods to customers. Fueled by the availability of cheap land and low wage labor, and facilitated by format evolution and information technology, the United States has sharply increased its consumption of this retailing service. Parts of the retailing sector have sought competitive advantage through low prices or convenience, whereas other parts offer specialized inventories, advice and luxurious surroundings. We are not aware that current price and output measures attempt to capture the way in which the quality of retailing service has changed over time. The productivity measure from the McKinsey Netherlands team was just a start, but it suggested that the impact of service quality differences may be quantitatively significant.

10.5 Public Transportation

Almost all of the McKinsey industry case studies have been of industries in the market portion of the economy—that is, where market-based

transactions and pricing typically prevail.[13] This was mainly by design; productivity comparisons are much more difficult where there is not a market valuation of output, and McKinsey's experience has traditionally been in the market economy. Recently, however, our European consulting practice has been developing more expertise in nontraditional areas like the provision of public services. Given the privatization and reform of public services that have taken place in Europe in recent years, our Netherlands office felt that these services might represent an important area of improvement potential for the Dutch economy. We selected public transportation as an industry in which output was relatively straightforward to measure yet had also been historically sheltered from market forces in the Netherlands.

10.5.1 Productivity Measurement

In the telecom case we defined output as including both the use of telecom services (call minutes) and the creation of the option to use telecom services (access lines). Public transportation is another industry in which the provision of an option to use services is an important component of output. Two half-full buses running a route every thirty minutes represent more output than a full bus running a route every hour. Even though the same number of people travel, those that traveled had more choices about when to travel, and thus experience a higher degree of convenience. This higher convenience is similar to that associated with having phone lines in every apartment (rather than a common phone in the hallway), more frequent flight connections due to a hub-and-spoke system, faster turn around times on a mortgage, or a conveniently located retail shop with exactly the fashions you want to buy.

Assigning a value to more convenient service is at least as difficult in public transportation as it is in other services. In retail, one can observe the margins paid in discounters and more focused specialty retailers that compete for the same customers and attribute the difference in margin to the extra value offered by the specialty store. But in public transportation more convenient and less convenient systems do not compete directly, and even when they do, fares are usually not as "market-determined" as are margins in retail.

As in the other industries (telecom, banking, and airlines) in which we have had difficulty deriving a market valuation of convenience, in public transportation we approached this issue from the input requirement side. In public transportation, labor input requirements are most directly deter-

13. All told, the McKinsey Global Institute has studied eight service industries (airlines, banking, film/TV/video, public transportation, restaurants, retail, software, and telecom) and twelve goods-producing industries (auto, auto parts, beer, computer hardware, construction, consumer electronics, furniture, metal working, processed food, semiconductors, soap and detergents, and steel). We have conducted country studies of Australia, Brazil, France, Germany, Japan, Korea, the Netherlands, and Sweden.

mined by the vehicle kilometers of service offered, not the passenger kilometers consumed. We collected data separately for four different modes of passenger transportation (long distance train, metro/subways, trams, and buses) in order to test for possible mix effects.

A difficulty in comparing productivity across countries is that city size and structure can influence the viability and efficiency of public transportation. In order to ensure that these differences were not driving the results, we compared five cities with similar populations, sizes of bus/tram/metro networks, and roughly similar population densities: Amsterdam, Rotterdam, Stuttgart, Zurich, and Stockholm. The city results broadly confirmed the results we obtained at the country level.

10.5.2 Productivity Results

Sweden had the highest productivity at the country level, and Stockholm had by far the highest productivity in our city comparison (table 10.5). The Netherlands placed fifth out of seven in our country comparison, ahead of only France and Germany, whereas Rotterdam and Amsterdam were third and fifth out of five in the city comparison. The Swedish productivity advantage was so large that despite higher service frequencies and lower usage, passenger kilometers per hour worked were also higher in Sweden (table 10.6).

Why was Sweden more productive? Two possibilities were that Sweden had a more productive mix of modes or that Sweden was obtaining higher labor productivity at the expense of low capital productivity. But Sweden had significantly higher productivity in both long-distance trains and in buses/trams/metros, and the productivity differences across buses/trams/metros were insignificant. Annual loaded hours per vehicle were roughly the same in Sweden and the Netherlands (18 percent higher and 13 percent lower in trains and buses/trams/metros in Sweden, respectively), suggesting that capital-labor substitution did not account for the differences.

The most important reason that labor productivity was higher in Sweden was simply that labor was more efficiently utilized. Driver, maintenance, and overhead labor were 30 to 50 percent lower per vehicle-hour in Stockholm than in Amsterdam or Rotterdam. In Amsterdam only 35 percent of paid driver time is actually spent driving, compared with 59 percent in Stockholm. In addition, average driving speeds are higher in Stockholm due to lower levels of congestion and higher penetration of high-speed rail.

Part of the difference in labor utilization is due to work rules. Work rules in public transportation are designed to provide security and protection for employees and passengers; but if applied too rigidly, these rules can have the side effect of lowering productivity. For example, in the Netherlands drivers are permitted a fifteen-minute break every four hours of driving to reduce fatigue and improve safety. When buses operate on hourly

Table 10.5 **International and City Comparison of Economic Performance in the Public Transportation Sector, 1995 (Index: NL = 100; and AM = 100)**

	Netherlands	France	Germany	Switzerland	United Kingdom	Japan	Sweden
Output	100	79	97	193	169	135	215
Labor productivity	100	61	87	113	125	144	188
Employment	100	131	111	171	136	93	115
	Amsterdam	Rotterdam	Stuttgart	Zurich	Stockholm		
Output	100	148	103	367	347		
Labor productivity	100	169	152	243	415		
Employment	100	87	68	151	84		

Source: McKinsey Global Institute and the Max Geldens Foundations for Societal Renewal (1997).

Table 10.6 **Alternative Productivity Comparisons between Country and City, 1995 (Index: NL = 100; and AM = 100)**

	Country Comparison		City Comparison		
	Netherlands	Sweden	Amsterdam	Rotterdam	Stockholm
Demand productivity					
Passenger kilometers per hour worked	100	130	100	100	261
Supply productivity					
Vehicle kilometers per hour worked	100	188	100	169	415
Vehicle utilization					
Passengers per vehicle	100	69	100	59	63

Source: McKinsey Global Institute and the Max Geldens Foundations for Societal Renewal (1997).

schedules, however, tight adherence to this rule can reduce labor usage significantly because the actual running time of the route can only be forty-five minutes. Otherwise, the fifteen-minute break every four hours would throw off the entire schedule. Allowing one five-minute break every hour would both provide more break time and increase productivity. In addition to this rule, Dutch public transportation is also subject to numerous other rules, mandating, for example, a five-minute boarding/mounting time, a twenty-minute turn-out time, a four-minute walking time after breaks, a second paid break after seven hours of accumulated working time, and in some locations (including Amsterdam) a fifteen- to twenty-minute end-of-the-line break. In Sweden, unions have allowed work rule revisions that enhance productivity in exchange for higher pay, and drivers have an incentive to contribute to higher utilization due to performance-based wages.

The emphasis on productivity in Sweden can be traced partly to increased market competition. Since 1989, local and regional bus, tram, and metro services have been exposed to competition through competitive bidding for three- to five-year concessions for parts of local/regional networks. The essential feature of this system is that local governments first decide on the service frequency desired, and then the operators compete on the cost to fulfill these service requirements. To minimize entry barriers in the operation of services, the local authorities have split the incumbent public transportation companies into two parts: One, which remains closely related to local government, is responsible for planning and designing the network, owns the rolling stock, and runs the bidding process for specific concessions. The other part of the incumbent companies was split into operational companies that compete for the concessions. Around 70 percent of bus, tram, and metro services are now managed this way in Sweden. Similar bidding systems have been put in place in other high productivity countries: the United Kingdom and Japan. Labor utilization and productivity have increased more rapidly in countries in which market competition has been introduced, and service frequencies have generally increased, not decreased as might have been feared. From 1980 to 1995, productivity increased 71 and 45 percent in the United Kingdom and Sweden, both of which introduced market competition, but increased only 20 and 12 percent in the Netherlands and France, which did not.

10.5.3 Lessons from the Public Transport Case

Public transport provides another example of a service industry that produces outputs that are difficult to value at market prices. In this industry, as in banking, telecom, and airlines, productivity measures can be developed by assuming that the economic value of a standardized output is related to the input required to produce it. What we learn from our public transportation case (and from banking, telecom, and airlines) is

that meaningful measures of cross-country output and productivity can be developed using relatively simple methodologies; that is, they result in measures of productivity that are consistent with what detailed knowledge of the production processes in the industry would lead one to expect.

10.6 Airlines

The airline industry is only a small part of the economy, but it is one that is studied intensively. It is an industry that provides some useful lessons about the effects of regulation and government ownership and the impact of changes in the quality of service. In particular, customer value in this industry is driven heavily by the ability of airlines to provide frequent service between any two destinations. In the United States this has meant that once the industry was deregulated, the major airlines soon found it essential to their competitive positions to develop the hub-and-spoke system.[14] This system results in some obvious congestion costs. As multiple flights arrive together at a hub, passengers find a congested terminal; ground personnel have to service the planes all at the same time; and air traffic controllers face take-off and landing delays. The European industry has only recently been deregulated, and indeed is still not fully competitive, so it has been slower to develop hub and spoke. But the same forces are now driving that industry also.

10.6.1 Productivity Measurement

The main international productivity comparison was between a sample of nine U.S. airlines (America West [HP], American Airlines [AA], Continental [CO], Delta [DL], Northwest [NW], Southwest [SW], Trans World [TW], United [UA], and US Airways [US]), and eight European airlines (Air France [AF], Alitalia [AZ], British Airways [BA], Iberia [IB], KLM [KL], Lufthansa [LH], SAS [Scandinavian; SK], and Swissair [SR]). As part of our studies of Brazil and Korea, we have also looked at productivity in the airlines of these two countries (Varig, VASP, Transbrasil, TAM, and Rio Sul [a subsidiary of Varig] for Brazil; and Korean Air and Asiana for Korea). We will discuss briefly the productivity findings for these industries also and how they differ from the United States and Europe.

We have estimated both labor and capital productivity, allowing an estimate of total factor productivity for the industries. For capital productivity we consider only the capital services from airplanes. The main reason for excluding IT and ground equipment is the difficulty in obtaining reliable numbers. Also, IT and ground equipment amount to only a small fraction

14. Even prior to deregulation there were elements of a hub-and-spoke system, as busy airports such as Atlanta or Chicago were frequently points of plane changes. However, regulation prevented the full development of the system with multiple hubs and complete spokes.

of the total physical capital stock of an airline (approximately 25 percent). We estimate the current market value of all the planes of the carriers included in this study and depreciate it over the remaining lifetime of the plane (maximum lifetime = 25 years). Leased aircraft are treated the same as purchased aircraft. The physical output for capital productivity is revenue passenger kilometers (RPK).

For labor productivity we followed a business activity approach that is natural for business consultants and is quite revealing in terms of diagnosing the reasons for the productivity differences, but that is not a standard approach among economists. Each functional group of employees was assigned an output measure reflecting the particular tasks they were engaged in, and the productivities of each employee group were weighted by labor shares, as in banking and telecom. For cabin attendants we used RPK; for ground handling and ticketing we used number of passengers flown; and for pilots and maintenance we used hours flown. Because the two products of "getting onto the plane" and "being flown *X* kilometers from point A to point B" are not separately priced, this approach (which is equivalent to weighting the functional outputs with unit labor requirements) provides an automatic adjustment for differences in stage length. In this way it is better than the traditional method of using only RPK, because this approach captures the fact that three 1,000-mile flights are typically priced higher, require more input, and provide more customer value than one 3,000-mile flight. We did check our results, however, using an overall measure of RPK per employee; and because average stage lengths are not very different across the countries we studied,[15] we obtained roughly similar results.

The labor productivity numbers are adjusted for differences in degree of third-party outsourcing of ground handling, ticket sales, and maintenance among the comparison countries. We also excluded cargo-only employers to be consistent with our output measures. The data used throughout this case study are drawn from international aviation databases and reports, including statistics provided by the International Air Transport Association (IATA) and the U.S. Department of Transportation (DOT). The information for Brazil was also obtained from the Departamento de Aviacao Civil (DAC).

10.6.2 Productivity Results

Table 10.7 shows the labor, capital, and TFP results for Europe, Brazil, and Korea, with the United States indexed to 100 for 1995. We found that the European airlines operated with about two-thirds the labor productiv-

15. Average stage lengths are much longer for airlines like Qantas and Singapore Airlines. This is one reason why studies that use only RPK as an output measure can yield misleadingly high relative productivity for these airlines.

Table 10.7 **Airline Productivity, 1995 (U.S. = 100)**

	TFP[a]	Labor Productivity	Capital Productivity
United States	100	100	100
Europe	75	66	92
Brazil	61	47	79
Korea		100	

Source: McKinsey Global Institute (1998a).

[a]TFP = (Labor Productivity)$^{0.5}$ × (capital productivity)$^{0.5}$

ity of the United States; Brazil had productivity less than half; but Korea had the same level of productivity as the United States. When MGI compared the United States to Europe in an earlier study, we found that European labor productivity was 79 percent of the U.S. amount. Moreover, the earlier study estimated productivity on a per-employee basis without adjustments for the differences in hours worked per full-time employee—an adjustment that would have raised European productivity. Although the sample of airlines was somewhat different in the earlier study, these figures suggest that Europe has fallen further behind the United States in the six-year interval after 1989.

The gaps in capital productivity, reflecting differences in RPK per plane, were much smaller, with the United States and Europe being essentially the same, and Brazil at almost 80 percent of the United States figure. Data on the stock of planes in Korea was not available. The European airlines are able to maintain about the same aircraft utilization as the U.S. airlines. The load factors are lower in Brazil. The TFP figures give equal weight to labor and capital, reflecting a rough estimate of the income shares in value added.

Figure 10.3 shows the breakdown of the labor productivity figure into the performance of the different groups of workers.

The most striking differences in operations occur in the ground personnel, including handling, maintenance, ticketing, sales and promotion, and other. As in the prior study, we found substantial overstaffing in these activities, reflecting the legacy of state-owned companies, inflexible workrules that reduce multitasking, and weak competitive pressure for cost reduction.[16] The airlines in Brazil have many of the characteristics of the European airlines, only more so. Varig was state owned until fairly recently and is now employee owned. It has received substantial tax breaks to maintain solvency. Aggressive competitors such as TAM are entering the domestic market, but slots and route allocations are controlled by the Air

16. British Airways, the European carrier that has been privatized and that competes against U.S. carriers, has productivity comparable to the U.S. airlines.

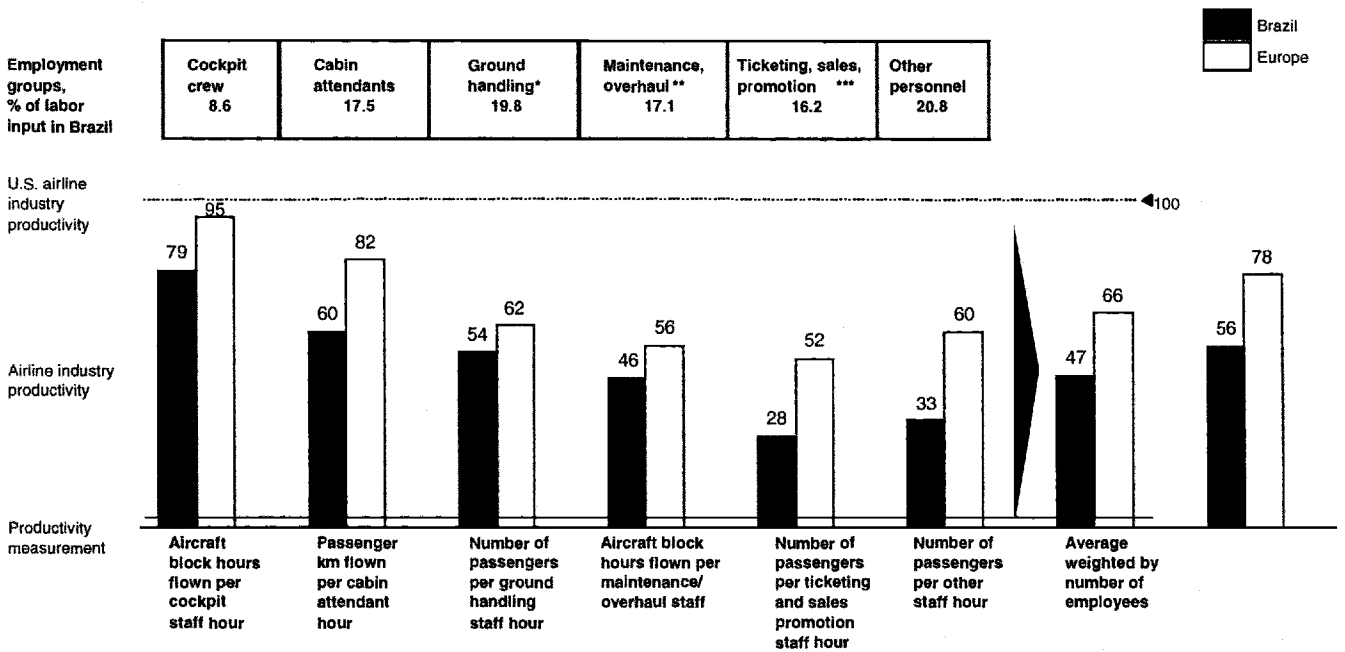

Fig. 10.3 Labor productivity levels in the airline industry, 1995 (U.S. productivity = 100)

Source: McKinsey Global Institute (1998a)

* Adjusted for 10% higher outsourcing to third parties for Brazil

** Adjusted for different degree of third party outsourcing

*** Adjusted for different percentage of tickets sold through airlines directly

Force. Brazil also suffers from IT inefficiencies. Importation of IT was prohibited for many years and remains costly. The Brazilian airlines in 1995 were behind Europe and the United States in using computers for scheduling and load management.

The differences in maintenance productivity reflect in part the impact of standardization. Having many planes of the same type (Southwest has all 737s) gives an advantage in maintenance. The U.S. airlines are able to reap economies of scale in maintenance, whereas Brazil has too few planes and too many types of planes for efficient operation.

We obtained less detailed information about Korea. Unlike many cases in Korea there is fierce price competition in the industry even though there are only two players. This has driven down ticket prices (average domestic air fares in 1995 were 13 cents per km in the United States, 19 cents in Brazil, and 8 cents in Korea) and promoted cost reduction and efficiency, at least in labor use. We have consistently found that intense price competition in an industry is the best way to force high productivity, but it is not always true that having two competing companies is sufficient, as apparently it is for Korean airlines.

10.6.3 The Impact of the Hub-and-Spoke System

Shortly after deregulation freed managers to restructure their networks, the U.S. industry moved to a hub-and-spoke network. By the mid-1980s, the hub-and-spoke system had already revolutionized the U.S. airline industry; and at the end of the decade, U.S. airlines operated about 30 hubs that, in terms of flight pattern and frequency, produce a different output than do nonhub airports.

The hub technology is distinguished from the operation of a large airport or homebase like Frankfurt or London-Heathrow by the coordination of the incoming and outgoing flights and the resulting flight pattern. Several times during the day, waves of flights come in and leave again about 60 to 90 minutes later. In between passengers from any arriving flight can transit to a connecting flight. In 1989 for instance, Northwest's hub in Minneapolis/St. Paul had about eight waves or banks per day with the biggest departure peaks at about 9:00 A.M., 1:00 P.M., and 6:00 P.M. to 7:00 P.M. The arrival peaks were roughly one hour earlier.

The transition created by the 1978 deregulation of the U.S. airline industry has sometimes been characterized as a shift from a linear system of direct nonstop and one-stop flights to a hub-and-spoke system involving connections at a set of central points where passengers are exchanged among flights. This is misleading because the frequency of passengers making connections barely changed, from 28 percent in 1978 to 32 percent in 1993 (Morrison and Winston 1995). Instead, what happened was that major carriers that had previously established hub operations serving a restricted set of spokes were allowed to increase the number of spokes

served from large hubs like Chicago and Atlanta greatly, whereas other carriers (some previously small regional airlines) established competing hubs that came to be dominated by a single carrier, sometimes as the result of a merger between two carriers previously serving the same hub. As the largest carrier in each hub city expanded into all the feasible spoke cities, competing airlines that previously had possessed protected rights on those routes were forced to abandon them. The result was the virtual elimination of interline connections, which in 1978 had represented 14 percent of all passengers and half of all connections. The remaining two-thirds of passengers who did not make connections in 1993 were served, as in 1978, by an abundance of nonstop flights on busy city-pair markets, both short-haul and transcontinental, and by nonstop flights on many new routes to spokes from the several new hubs.

The transformation of the U.S. airline industry to the hub-and-spoke system was driven by demand and revenues, rather than by cost. Given the traffic dispersion in the United States and given customers' desires for frequency between origination and destination cities, the hub-and-spoke technology is the only economical way to offer fast and frequent air transportation at times convenient to most of the travelers. Consequently, the carrier that could offer an advanced hub-and-spoke network could attract more customers and gain a competitive advantage over its competitors.

Whereas the hub-and-spoke system has increased service quality and variety for most of the customers (particularly higher frequency of flights), we think that this innovation has generally had a negative impact on measured labor and capital productivity, given that we are not able to measure the added customer value as an output. The productivity price for the hub-and-spoke technology is extremely peak-driven operations at the hub, with unfavorable implications for labor and capital utilization. This means that our physical output measures are missing an important qualitative aspect of the U.S. airline industry—namely the overall network performance, or the ability to provide frequent services to hundreds of destinations in a reasonable time. To the benefit of the consumer, this output quality is provided at the expense of higher factor inputs, without being captured in our productivity measure.

10.6.4 Lessons from the Airline Study

The main implication for productivity measurement from our case study is that improvements in the quality of service should be considered in a correct measure of output for this industry. Airlines advertise the quality of their food and the friendliness of their personnel (arguably attributes of service that have deteriorated since deregulation), but it is MGI's conclusion from working with many airlines around the world that frequency and reliability of service are the key factors in service quality, particularly among business travelers. This means that the industry developed a major

innovation in the 1980s in the hub-and-spoke system after deregulation, the benefits of which were not captured in productivity measures. Quantitative estimates of the productivity penalty associated with hub-and-spoke or of the increased customer value from increased service would have been valuable. Since the development of this system has been the principal structural change occurring in this industry since 1980, it is likely that its impact on true productivity has been substantial.

This conclusion bears on the debate about the impact of IT. As we have seen from the experience of Brazil, operating a point-to-point airline with inadequate IT exacts a productivity penalty. Operating a hub-and-spoke system without IT would be impossible. The share of IT hardware in the total capital stock of the industry is small. The impact of the technological change facilitated by IT has been large.

10.7 Summary and Conclusions

The first step in correctly measuring the output of an industry is to understand that industry: how it is changing and how this change is being driven by consumer demand. Although we have used relatively simple measures of output in the service industries we have studied, we have generally been able to satisfy ourselves that the resulting productivity comparisons make sense in terms of what is going on at the production process level and how the industries are evolving.

In banking, we found that an index of output, based on the three main functions of retail banks, performed pretty well as an indicator of relative performance. An important lesson from this work is that shifting to electronic funds transfers has a very large impact on productivity. In general, MGI has found that increased competition improves productivity, but banking in the Netherlands provides one example where a relatively concentrated industry was able to make the shift in technology more easily. Hopefully, government coordination through Federal Reserve actions will accomplish the same goal in the United States while maintaining competitive intensity.

Despite the useful lesson from the simple output measure in banking, the case illustrates the limitations of our measurement. An important change in the mortgage industry has been the introduction of streamlined processes and computerized credit assessments. This has reduced the intermediation margin and speeded up the approval of loans. Neither of these changes in cost or in quality is captured in our output measure or in others that we know of. Further, the growth of the securities industry has been very rapid. Much of the activity consists of selling services and giving investment advice, and the intrinsic value of both activities is hard to value. An additional issue has arisen when we have extended our analysis to countries such as Korea, whose banking system allocated loans without

adequate concern for the returns and found itself with a seriously bad debt problem even prior to the current crisis. Measuring banking output on the wholesale side would pose new challenges.

The banking case gave clear evidence of the importance of IT on productivity. ATMs, check processing equipment, use of terminals to access information, call centers, loan processing, and of course electronic funds transfers are all important in this industry.

Looking forward, we anticipate substantial changes in the retail banking industry. It is possible that the main functions currently carried out by banks will migrate to other institutions. To an extent this has already happened, as specialized lending institutions have taken over large parts of the credit card and mortgage markets. Retailers already provide ATMs, credit cards, and cash-back services and may expand to offer most or all of the current services provided by banks. Unless measurement approaches evolve with the industry, errors may worsen in the future.

In telecom there are three main attributes of service in developed countries: the access line, the call minutes, and the mobility of the call. We were able to capture two of these. Clearly, as Hausman (1997) has shown, it would be possible to capture the value of the third element. We have ourselves explored a hedonic or conjoint approach to the problem in which we would be able to price mobility, given the alternate pricing schemes in effect in different countries, but we did not complete the task. In the developing country context, problems such as noise on the line, lack of call completion, and cutoffs become important elements in service quality. They remind us that over the long term, service quality has greatly increased in the United States.

Going forward, we expect that new technology will allow major improvements in service variety and quality. High-speed data transmission will allow new services to be provided. It is likely that phone service with picture as well as sound will improve in quality and become more common. The phone companies will soon allow people to have a single number and have calls that track the location of the person. (Hopefully one will be able to switch such services off as well as on.)

Information technology, notably digital switching, bill processing, and cellular service, has been important to productivity in this industry. It will be central to the changes in service going forward.

Retail has been the industry where measurement has been most difficult for our studies to resolve satisfactorily. This is one of the largest industries in the United States. It has changed dramatically over time, and employment has grown. It is a sector where consumers clearly value its output. The evolution of retailing formats from traditional stores to discounters, category killers, and high-service specialty stores has transformed the industry; indeed, it has transformed the cities and suburbs. Given the difficulties we have experienced in our cross-country comparisons, we are in

no position to preach to the statistical agencies about how to solve the measurement problem. However, we know there is a problem that needs to be solved.

The continuing trend in the industry is that discounters invade the product and service lines that are started by innovations among the specialty players. This lowers prices to consumers and passes back to them the benefits of innovation, a process familiar in all industries. Without new measurement methods, however, productivity growth will continue to be understated.

Information technology has played an important role in this industry in allowing retailers to eliminate the wholesale function and work directly with manufacturers. IT speeds checkout and tracks inventory. It allows companies to track demand and focus on the items consumers want.

Adequate tracking and benchmarking of public services can provide a spur to improved performance in the public sector. The Netherlands study made benchmark comparisons of public transportation services productivity, revealing the very high price paid by riders in Amsterdam for the work rules and restrictions in place in that city. Given the size and importance of the public sector, such benchmarking studies have tremendous potential importance.

Recent improvements in BLS tracking of discount fares have improved the measurement of prices and hence real output in the airline industry. Now that the hub-and-spoke system is firmly in place, changes in service frequency and reliability may or may not be major factors in service quality change going forward. In order to understand the historical development of productivity in the industry, however, it would be essential to factor in the effect of this innovation.

This industry has been adjusting to the impact of deregulation over the past several years. New entry has changed the competitive intensity and forced major restructuring. That process continues and it may or may not induce new innovations in service quality going forward.

Information technology was an important element in overall productivity for the industry. It allows for more efficient use of resources and means that companies can operate with the increased complexity of hub and spoke.

10.7.1 Common Themes

For the purposes of international comparisons, we have concluded that simple physical output measures work surprisingly well. These output measures can be related to the inputs used in their delivery, as we saw in banking, telecom, and airlines. We can make adjustments to the basic measures for quality differences, or we can ignore those that do not seem to make a substantive difference to the overall conclusions we reach about the causes of productivity differences. The one case where a simple physi-

cal measure was inadequate—retailing—is the case that was hard to solve any other way.

Analyses of these case study industries suggest that the simple output measures will be much less adequate for productivity growth over time. The knowledge that MGI can offer in this area is to point to the most important drivers of industry change. If the data collection techniques in use are not capturing the impact of industry evolution, then they are likely to be substantially in error. We have seen plenty of evidence of the importance of quality change and convenience.

A possible general approach to improving the measurement of these convenience and quality changes is to measure and value convenience-related customer time savings and to use hedonics to estimate market valuations of service quality differences where possible. Developing exact procedures will obviously require further study, but our cross-country findings suggest that quality and convenience are potentially very important sources of error.

All of the cases, with the possible exception of public transportation, showed that IT was a vital component of the industry business systems. Based on what we learned of these industries, the value of computer hardware was only a small component of total capital and, on a growth accounting basis, IT would make only a small contribution to labor productivity. Nevertheless, we found that the current business systems in these industries would be impossible without IT. A substantial increment to productivity is associated with its use. Brazil, which had specific regulations about IT purchases, and had productivity limited by inadequate IT in airlines and banks illustrates this point.

References

Council of Economic Advisers. 2001. *Economic report of the president.* Washington, D.C.: GPO.

Hausman, Jerry A. 1997. Valuing the effect of regulation on new services in telecommunications. *Brookings Papers on Economic Activity: Microeconomics, 1997,* 1–54.

McKinsey Global Institute and the Max Geldens Foundations for Societal Renewal. 1997. *Boosting Dutch economic performance.* Amsterdam: McKinsey & Co.

McKinsey Global Institute. 1998a. *Productivity—The key to an accelerated development path for Brazil.* Washington, D.C.: McKinsey & Company.

McKinsey Global Institute. 1998b. *Driving productivity and growth in the U.K. economy.* Washington, D.C.: McKinsey & Company.

Morrison, Steven, and Clifford Winston. 1995. *The evolution of the airline industry.* Washington, D.C.: Brookings Institution.

Comment Robert J. Gordon

The Baily-Zitzewitz paper continues the tradition of the McKinsey Global Institute in displaying admirable creativity and imagination to overcome seemingly intractable obstacles to making valid cross-country comparisons of levels of productivity. We emerge from the analysis with both new measures of the extent of productivity differences across countries in five industries and considerable understanding of the underlying sources of these differences. In this comment I begin by extracting from the paper some of the most interesting generalizations about the sources of cross-country productivity differences and subsequently discuss what we learn (and do not learn) about the "lessons for measurement" previewed in the paper's title. In the final section of this comment I offer a few thoughts about the treatment of airlines. To preview what comes later, my main quibble about the paper is that the lessons for measurement are directed more at the consulting community than at the practitioners of productivity measurement in national statistical offices; we learn less than one might hope about how some of the insights from the industry studies can be applied to the measurement of the time series evolution of productivity, either at the level of these five industries or for the economy as a whole.

Generalizations about the Sources of High Productivity

Perhaps the most striking of the paper's conclusions is that there is no single national model for achieving the highest productivity in every industry. Unlike the earlier McKinsey Global Institute study of the service sector (1992), in which the United States seemed to emerge as the clear leader in most industries, here the leadership role is shared, with the U.S. clearly ahead only in telecom and airlines, the Netherlands in banking, Sweden in public transport, and an ambiguous verdict in retailing (subsequently we return to the measurement issues specific to retailing). Another common theme that runs through the Baily-Zitzewitz paper is American exceptionalism, which has mixed implications for productivity, from backwardness about electronic payments in the banking industry; to admirably high usage of telephone lines; to a retailing system driven by low-cost land and lenient land-use laws; to public transit that is so nonexistent that it is not worthy of study in the paper; to a set of large airlines benefitting from economies of scale that in part reflect national geography, dispersed families, a common language, and lenient regulation of mergers.

Implicit in the paper's case studies is the generalization that government policy is capable of both harm and good. Some of the harm is created by

Robert J. Gordon is the Stanley G. Harris Professor in the Social Sciences at Northwestern University, a research associate of the National Bureau of Economic Research, and a research fellow of the Center for Economic Policy Research.

that age-old enemy of efficiency, namely high tariffs, evident in the paper's emphasis on the distorted prices of imported capital goods that hinder the development of the Brazilian telecom industry. Work rules in public transportation are another government-imposed source of inefficiency, as in the example of inefficient Dutch public transport contrasted with Sweden's. Shop closing hours and zoning restrictions are among the regulations cited by Baily and Zitzewitz to explain the more advanced development of mass-merchandising retail formats in the United States compared with Europe. Labor market regulations restricting the development of part-time employment interact with shop closing hours to reduce the convenience (and consumer surplus) enjoyed by European consumers compared to their American counterparts.

However, there are important measurement and policy issues that are implicit in the discussion of retailing but that never rise to the surface in the Baily-Zitzewitz presentation. Mass merchandisers achieve higher gross margins and sales per employee than do older formats and help the United States achieve its productivity advantage in retailing, but this has been achieved by a systematic and controversial set of U.S. government policies that has encouraged the development of metropolitan areas that by European standards have very low population densities. The most important policies that have steered the United States resolutely toward low density and high dispersal are the massive investment in interstate highways, which in most metropolitan areas have no tolls or user charges; the starvation of investment in public transport; the tax deductibility of mortgage interest payments, which encourages overinvestment in houses and large lots; and the dispersed governmental system, which permits local zoning regulations that restrict the minimum lot size in many suburban jurisdictions. Low density imposes numerous costs on American society that are hidden in standard productivity measures, including excessive investment in automobiles and highways that consume resources to an extent that is unnecessary in European cities, and the time and aggravation costs of suburban road congestion and long travel times from home to work. Thus the criticism of European zoning regulations for inhibiting the development of modern retailing formats implicit in the Baily-Zitzewitz paper must be set against the costs of low-density suburban development, which remain unmeasured either by the authors or in official measures of productivity.

The contrast between the European system and American system has increasingly attracted the interest of academic economists in recent years (see, for example, Gordon 1997). Low minimum wages and a relatively light burden of social security taxes in the United States greatly boost the wage for the marginal unskilled American employee, especially in the retail sector, and lead to rapid growth of employment in types of jobs that scarcely exist in some European countries, particularly grocery baggers

(mentioned by Baily and Zitzewitz), buspeople in restaurants, parking lot cashiers, and valet parking attendants. Thus in the U.S. retail sector two opposing forces affect productivity in retailing compared with Europe: The low-density suburbs and lax regulations on land use encourage the rapid development of mass merchandising and out-of-town specialty formats, boosting the level of productivity, whereas the heavy use of unskilled low-wage workers for grocery bagging and other tasks reduces productivity (while creating unmeasured increases in consumer convenience).

Far from causing universal harm to productivity, government policies can actually be the source of good. The much more advanced use of electronic transfers by the European banking system, and the quaint backwardness of the U.S. banking system, is traced by the authors partly to the role of the Netherlands Postbank in taking the lead in electronic payments and forcing privately owned banks to follow. Also important in explaining U.S. backwardness is the set of government regulations that until recently prevented interstate banking and created an inefficient system of thousands of unit banks. Even to this day, a wire transfer into or out of my own personal bank account is a rare event, almost always involving a payment coming from a European source, and the instructions for making a wire transfer to my account are so complex that seven lines of text are required.[1] Another lesson for policymakers, especially in the U.S. Department of Justice, is that in some cases a high level of concentration encourages efficiency by reducing administrative overhead and encouraging innovation, whereas low level of concentration can lead to inefficiency. Although primarily implied by the Baily-Zitzewitz study of banking, this conclusion is true for the European airline industry, where fleet sizes of particular aircraft types are too small to benefit from economies of scale in aircraft maintenance.

Measurement Issues

The Baily-Zitzewitz paper grapples imaginatively with the issue of measuring output in industries with multiple unpriced outputs, such as banking and telecom. They combine volume measures of such functions as telephone access lines and call units, or in the case of banking, processing

1. The reluctance of U.S. businesses to use electronic transfers extends far beyond the banking system itself. No credit card or other national merchant that sends me bills includes codes on the bill for electronic payment. For years I used an electronic bill-paying system run by Fidelity Investments, which allowed me to avoid writing checks but which required Fidelity to print computer-written checks and physically mail them to merchants. I abandoned this system because the mailing addresses and postal box numbers of the merchants changed so rapidly that keeping Fidelity informed of the changes became a headache compared to the old-fashioned method of sticking a check in a window envelope with a preprinted payment stub. Further, at least half my checks are written in idiosyncratic transactions that would not be worth setting up merchant numbers and account numbers for simple one-time payments to local merchants or workmen (all of whom would have electronic account numbers printed on their invoices if located in the Netherlands).

payment transactions, maintaining deposit accounts, and lending. Weights to combine the separate types of outputs are based on estimates of labor requirements rather than on costs or revenue shares, and the authors appeal to lack of data rather than theoretical purity to justify their choice. In some industries, especially airlines, there is enough data to run a regression to separate the per-passenger part of the fare from the mileage component, but the drastic change in the weight of these components after U.S. deregulation (with the per-passenger charge rising to reflect cost and the per-mile component declining substantially) reminds us that revenue weights in some industries and eras may be a reflection of regulations more than the basic economics of the industry.

We can assess the methodology in the Baily-Zitzewitz paper both within its own context, providing a consultant's expertise on the sources of productivity differences across countries, and on the more ambitious goal set forth in the paper's introduction, "to ask whether the results of these international comparisons cast any light on the problem of service sector output measurement." I have few objections to the case studies themselves. A purist might object that adjustments are introduced in an ad hoc way, as in the 5 percent adjustment for the lower quality of telecom service in Brazil. Symmetry would require a substantial adjustment to the productivity of European airlines for the higher labor requirements of ground handling caused by the costly necessity of separating domestic and international operations, and for additional expenses of remote ticket offices required by different languages and the more complex ticketing requirements of international tickets. I find adjustments in some industries and not others and wonder whether the results would be affected much (if at all) if the adjustments were equally thoroughgoing in each industry.

The paper is less successful in its broader ambition of providing lessons for the measurement of productivity in the services sector at the level of national statistical agencies. The central problem is that the basic mission of statistical agencies is to measure the level of labor productivity and multifactor productivity for their own economy over a time span consisting of many years and decades. Yet the paper never delves into the implications of its own analysis for the growth rate of productivity over time. For instance, the authors state that "even the BLS [U.S. Bureau of Labor Statistics] estimates may be missing changes in the true output of the sector," yet no attempt is made to substitute the authors' methodology in banking for that of the BLS or to measure the difference made by the two methodologies over, say, the last decade.

The authors' lack of interest in time series issues extends even to their unwillingness to use past evidence gathered in McKinsey (1992) to measure changes over time in the industries that both studies looked at: namely, banking, telecom, retailing, and airlines. The authors missed the opportunity to take their current weights and their measures of outputs

and inputs applied equally to the data used in the 1992 study and the current study, and to calculate annualized growth rates in productivity for the years between the two sets of data, and then to have compared and analyzed differences in these growth rates from those computed in official data for the same industries, for example, by the BLS and similar sources for other countries. For instance, how rapidly would their measure of telecom productivity of access lines and call units increase compared to the current BLS estimate that output per hour in the U.S. telecom industry grew at a rate of 5.4 percent per annum and the U.S. commercial banking industry at a rate of 3.1 percent per annum over the time interval 1987–96?

Another measurement issue closely related to the Baily-Zitzewitz analysis of retailing emerges in the Boskin commission report on the U.S. Consumer Price Index (CPI), the so-called output substitution bias. There are two extreme views regarding the difference in retail prices across retailing formats. At one extreme, the official view of the CPI is that any difference in price between, say, a ma-and-pa neighborhood drug store and Wal-Mart is entirely reflected in a difference in service. If the arrival of a new Wal-Mart in town causes the price of toothpaste to drop from $2.59 at the old-format store to $1.49 at the Wal-Mart, the entire price difference of $1.10 must be due to the superior service and location of the old-format store. This is implicitly the view of Baily and Zitzewitz, who state, "We believe that the format bridge approach is an appealing way to approach the problem. . . . We have tried to compare like retailing service to like retailing service."

At the opposite extreme is the view that the entire price decline of $1.10 in this example represents a price decline, with no difference in quality, and that the CPI is biased upward for ignoring this decline in price and attributing the entire difference in price to a difference in service quality. The Boskin Commission report seemed to adopt this position, although when pressed, most commission members including myself would assume that the truth lies somewhere between the two positions rather at the lower extreme of ignoring any difference in service quality.

Here, however, I would like to differ with both the CPI and the Baily-Zitzewitz paper. There are three persuasive reasons to believe that consumers view most of the price reduction available "when Wal-Mart comes to town" as a true reduction in price rather than as a reduction in service quality. By far the most important reason is that consumers have been voting with their feet and their autos, flocking to Wal-Mart. The market share of discount department stores rose from 44 percent to 68 percent between 1988 and 1998 as consumers chased the lower prices. If lower prices had been completely offset by lower service quality, there would have been no shift in market shares. Second, there is no difference in service quality between the newer-format mass merchandise discount stores (Wal-Mart, Target, Kmart) and the older-format department stores that cater to the same customers (Sears and Montgomery Ward). All these

stores rely on self-service, so there is no reduction in service quality when a customer shifts from old format to new format. The basic reason that the new-format stores charge lower prices is that they are more efficient, with innovative electronic inventory management systems pioneered by Wal-Mart and Target stores. Faced with this competition and fleeing customers, Montgomery Ward and Woolworths went bankrupt, further evidence that their price differential was not supported by superior service. Third, new-format stores may actually provide *superior* service. A front-page article in the *Wall Street Journal* (24 November 1999) traced part of the customer preference for Wal-Mart compared with Sears to the availability of shopping carts at the former but not the latter, and the ability to do all checking out on a single stop at exit checkout aisles in the new-format stores, instead of at separate islands in each department. This is reminiscent of the process by which grocery supermarkets made individual butcher and produce shops obsolete by combining lower prices with single rather than multiple checkouts.

When discussing upward bias in indexes of consumer prices, by far the most elusive area is the consumers' surplus provided by new products. At one point Baily and Zitzewitz allude to several ingredients that would help in the construction of a measure of the consumers' surplus provided by telephone communication, including the fact that rights to an access line in Brazil were trading at $5,000 in a country with only 22 percent of the value of U.S. real GDP per capita, and that U.K. customers demand over 1,300 annual minutes per capita at a price of 5 cents per minute, indicating that U.S. customers enjoying free local calls are enjoying a substantial consumers' surplus. The frustrated reader vainly hopes that the authors will work out these implicit calculations and express them as a ratio to total consumption or real GDP, but the reader receives no help from the authors and wonders whether 1,300 annual minutes $\times$ $0.05 (i.e., a mere $65) actually warrants the adjective "large." Also, we do not learn whether the total cost of phone calls in the United States is less, or whether there is just a redistribution between light and heavy users implied by a higher access charge and lower unit cost in the United States as contrasted with the United Kingdom.[2]

The Airline Industry

Because airline management is my particular area of industry expertise, I will conclude with a few comments on the paper's treatment of airlines. The current paper is more sophisticated than the related study of airlines

2. Also, the presumption that local calls are typically free in the U.S. may be an obsolete observation. Local phone service in the United States has become inordinately confusing, with multiple telephone bills for the same phone line; and on inspection of my phone bill for October 1999, I discovered that for my main home phone line I had no free call privileges but instead was charged at a rate that averaged $0.0338 per minute, hardly different from the author's quoted rate for the United Kingdom several years ago.

in the McKinsey (1992) volume but still requires a few comments and corrections.[3] The authors distinguish the U.S. hub technology, which they agree started before deregulation in 1978 but further developed (by opening new hubs and filling in missing spokes in old hubs) by the mid-1980s, from a "large airport like Frankfurt or London-Heathrow." They seem to imply that the major home bases of the large European airlines are something different than the hub-and-spoke system operated by the U.S. airlines. The only sense in which their comment is correct is that London-Heathrow is so constrained by slot limitations that of necessity arrivals and departures are spread evenly over the day and do not exhibit the peaking of scheduled banks of flights characteristic of U.S. hubs, which (other than Chicago O'Hare) are not subject to slot controls. But the other large European airports were operating hub-and-spoke systems by the early 1970s that were much more developed in the United States simply because by definition each large European airline had landing rights between its home base and each outlying spoke, whether in Europe, the United States, Asia, Africa, or Latin America. Because the only two U.S. airlines offering substantial service to Europe at that time (TWA and PanAm) funneled their passengers through New York's Kennedy airport, for many residents of U.S. cities like Chicago or San Francisco, the best and fastest way to get to most cities in Europe was to fly on a foreign airline to its hub, whether London, Amsterdam, Frankfurt, Zurich, or Paris, and continue on that same airline to the spoke city (Lisbon, Rome, etc). Today Amsterdam, Frankfurt, and Paris De Gaulle are large hubs with all the characteristics of the major U.S. airports that Baily and Zitzewitz describe, with banks or waves of flights arriving periodically, and idle periods in between. Amsterdam and De Gaulle each have six banks of flights daily, fewer than O'Hare, Atlanta, Minneapolis, and other large U.S. hubs; and the peaking of flight operations is evident in Amsterdam between 8:30 A.M. and 10:30 A.M., when seventy-five jets depart, or in Frankfurt from 12:30 P.M. to 2:00 P.M., when a like number of jet departures occur.

There is no argument with the authors' conclusion that the development of U.S. hubs was driven by demand and revenues and is doubtless cost-inefficient by peaking baggage sorting and other ground operations and leaving idle periods between. The cost inefficiency is particularly evident at smaller hubs like Northwest Airlines' operation at Memphis, with just three banks per day; and so costly are small hubs with infrequent banks that American Airlines closed down three of them (San Jose, Nashville, and Raleigh-Durham) in the early 1990s. My quibble is that this does not represent a difference between the United States and Europe, either for

3. I was a discussant of the published article drawn from the Baily (1992) study. The discussion of airlines in the current study adopts many of the suggestions made in my earlier comment, particularly about the nature of the transition of the U.S. route structure after deregulation.

the larger European hubs discussed here or the smaller hubs like Brussels, Zurich, Copenhagen, and Milan.

The major omission in the airline case study is the failure to recognize that at least some of the higher labor requirements of European airlines are due to the inherently more international nature of their operations. The typical European airline has a far greater percentage of its flights that cross borders to another country than does the typical U.S. airline and thus faces physical barriers within airports that require duplicate staff for domestic and international flights. The Schengen agreement that has lifted passport controls among many European countries has, if anything, made the situation more complex and labor intensive. The Frankfurt airport was required to build a whole new set of third-floor lounges so that it could service the same gates from a second-floor lounge if extra-Schengen and a third-floor lounge if intra-Schengen; and passengers are frequently required to travel by bus to aircrafts that are parked at jetways and could have been reached by foot through the regular concourses but for the rules that require bypassing passport control stations. The extra control personnel and bus drivers may in some cases be on the payroll of the airlines and others on the payroll of airports, but they doubtless account for some of the European airlines' inefficient labor use.

Finally, part of the European airlines' inefficiency is due to differences in language across the continent. On one visit I found that Lufthansa had two people working in Florence, Italy, in a ticket office servicing one flight per day. Not surprisingly, Baily and Zitzewitz show that Europe's worst functional productivity ratio is in ticket and sales personnel. Why doesn't United Airlines, which serves places like Burlington, Vermont, with two or three flights per day, have a ticket sales office in those cities? Most of the difference must be due to language. German travelers in far-off Florence want the security blanket of a Lufthansa office where they can speak German if they want to change their plans or reconfirm their reservations. Airlines in the United States do all this with monolingual continent-wide toll-free numbers.

These observations do not imply that the paper made mistakes in its analysis of airlines, but rather that it omitted some of the sources of the U.S. productivity advantage. Just as it is a time-honored principle of economics that tariffs breed inefficiency, so too it is a time-honored source of the century-long American productivity advantage that numerous industries, especially in transportation, benefit from the economies of scale implicit in a continent-wide zone free of barriers to the movement of people or goods and able to communicate in a common language. Perhaps the greatest surprise suggested by the current paper is that Americans remain so steadfastly backward in writing checks instead of carrying out their financial transactions electronically. The second greatest surprise, not discussed explicitly in the paper, is that Europe has raced far ahead of the

United States in mobile phone technology and usage by agreeing to a single standard, instead of the four competing standards in the U.S. cellular phone industry. Sometimes, it makes sense for the government to impose industry concentration and uniformity.

References

Baily, Martin Neil. 1993. Competition, regulation, and efficiency in service industries. *Brookings Papers on Economic Activity, Microeconomics:* 71–130.

Gordon, Robert J. 1997. Is there a tradeoff between unemployment and productivity growth? In *Unemployment policy: Government options for the labour market,* ed. D. Snower and G. de la Dehesa, 433–63. Cambridge: Cambridge University Press.

McKinsey Global Institute. 1992. *Service sector productivity.* Washington, D.C.: McKinsey & Co.