

# CESifo DICE REPORT

## Journal for Institutional Comparisons

VOLUME 3, No. 4

WINTER 2005

---

### Forum

#### RAILWAY (DE-)REGULATION

Marc Ivaldi  
Fumitoshi Mizutani  
Chris Bolt  
Günter Knieps  
Jan Scherp

---

### Research Reports

GENERATIONAL INCOME MOBILITY  
LABOUR MARKET INSTITUTIONS  
TAX BURDEN OF COMPANIES

Miles Corak  
Wolfgang Ochel  
Michael Overesch

---

### Reform Models

EDUCATION POLICY IN THE UK

Stephen Machin and  
Anna Vignoles

---


### Database

WOMEN'S ATTITUDES TOWARDS FAMILY  
DEVELOPMENT AID  
DOING BUSINESS  
MORTGAGE DEBT  
COLLECTIVE BARGAINING  
TRADE RESTRICTIONS

---

### News

NEW AT DICE DATABASE,  
CONFERENCES, VALUES SURVEYS



**CESifo**  **DICE** Search














**Simple** **Advanced** **Recent Files** Survey Help

Search  Show results:  in  format, sorted by

Text

Text search initiates searches in document texts as well as the attributes title and comments.

Select the institutional sector or subsector (folder) you want to search. Click on the   images to expand / collapse folders. Selecting no folders initiates searches in all institutional sectors and subsectors.

-  Basic Country Characteristics
-  Capital Market
-  Education
-  Enterprise Environment
-  Experts' Assessments of Governance Characteristics
-  Health
-  Labour Market
-  Natural Environment
-  Pensions
-  Public Finance
-  Social Policy
-  \_Archives
-  \_Miscellaneous

*Visit the DICE Database at  
[www.cesifo.de/DICE](http://www.cesifo.de/DICE)*

CESifo DICE Report ISSN 1612-0663

A quarterly journal for institutional comparisons

Publisher and distributor: Ifo Institute for Economic Research e.V.

Poschingerstr. 5, D-81679 Munich, Germany

Telephone ++49 89 9224-0, Telefax ++49 89 9224-1462, e-mail [ifo@ifo.de](mailto:ifo@ifo.de)

Annual subscription rate: €50.00

Editors: Wolfgang Ochel ([ochel@ifo.de](mailto:ochel@ifo.de)), Rigmar Osterkamp ([osterkamp@ifo.de](mailto:osterkamp@ifo.de))

Editor of this issue: Wolfgang Ochel

Copy editing: Anne Heritage, Paul Kremmel

Reproduction permitted only if source is stated and copy is sent to the Ifo Institute.

**DICE Database: [www.cesifo.de/DICE](http://www.cesifo.de/DICE)**

**Forum****RAILWAY (DE-)REGULATION****Restructuring Railways: In Theory and Practice***Marc Ivaldi*

3

**Regulation and Deregulation in the Japanese Rail Industry***Fumitoshi Mizutani*

10

**Railway Privatisation and Regulation in Great Britain***Chris Bolt*

16

**Railway (De-)regulation in Germany***Günter Knieps*

21

**Railway (De-)regulation in EU Member States and the Future of European Rail***Jan Scherp*

26

**Research Reports****Inequality across the Generations in North America and Europe***Miles Corak*

34

**Concepts and Measurement of Labour Market Institutions***Wolfgang Ochel*

40

**The Effective Tax Burden of Companies in Europe***Michael Overesch*

56

**Reform Models****Education Policy in the UK***Stephen Machin and Anna Vignoles*

64

**Database****Women's Attitudes towards Family and Gender Roles**

75

**Trends in Development Aid**

78

**Where is Doing Business Easy?**

81

**Mortgage Debt**

84

**National Collective Bargaining Systems in Europe**

86

**Trade Restrictions**

89

**News****New at DICE Database, Conferences, Values Surveys**

92

# RAILWAY (DE-)REGULATION

## RESTRUCTURING RAILWAYS: IN THEORY AND PRACTICE

MARCIVALDI\*

### Introduction

After years of permanent decline, the European railways industry is experiencing changes. A clear sign of these changes is found in the coexistence of the old and the new at the present time, which yields a highly contrasted picture. Here the passenger is wondering if the nightmarish train in which he/she finally finds a relatively clean seat belongs to the same world of this millennium; there he/she is cocooned while moved at 300 kph (and soon even more) on a high speed train between Paris and Brussels. Whether these changes are significant, and whether the policies underlying them are relevant for this industry to take on a new lease of life, to be part of a sustainable growth in Europe and, in some ways, to recover the glory attached to its role in the first half of the 20th century, motivates a lively debate in Europe and elsewhere.

This article is aimed at shedding some light on some of the main economic issues that are on the agenda of this evolving industry.<sup>1</sup>

The last three decades have seen a dramatic decline of EU railways while the transport industry in Europe as a whole has grown steadily at a 3 percent average rate. Rail freight transport has lost two thirds of its market share to the benefit of road and maritime freight transport. Over the period 1970-2000, its market share has decreased to 8 percent from 21 percent. Although of a smaller size, the loss is also sig-

nificant for rail passenger transport, whose market share has vanished to 6 percent from 10 percent (Di Pietrantonio and Pelkmans (2004) and Nash, Matthews and Shires (2004)).

The European case contrasts with the state of the US railways industry. The restructuring of US freight markets undertaken in the early 1980s has allowed rail industry to dramatically improve its efficiency-reducing both its workforce and its track mileage by 50 percent. In the face of competition between truck and rail and between integrated firms, the rail freight industry has maintained its level of activity by focusing especially on bulk and intermodal shipments.<sup>2</sup> The rail share of the freight market on a tonnage basis was 11.8 percent (compared to 54.5 percent for truck) in 1993 and 12.2 percent (compared to 58.2 percent for truck) in 2002. It is also important to note that during the period 1980-2005, the US rail industry experienced a significant merger wave and achieved a high level of concentration. Despite this dramatic industry consolidation, the consumer surplus in US rail freight markets increased by about 30 percent between 1986 and 2001, suggesting that to date the trade-off between merger-specific efficiency gains and merger-related increased market power has favoured rail customers (Ivaldi and McCullough 2005).

In worldwide perspective, however, the railway industry faces strong challenges from other modes of transportation. All attempts to restructure this industry focus on two main questions: What is the best structure for this industry? Is competition feasible and what are the conditions for efficient entry? We provide some answers to these questions.

### Vertical disintegration

In all network industries, vertical disintegration is a key tool for reforming old utilities. In railways, vertical disintegration is to be viewed as the separation of



\* Marc Ivaldi is Professor of Economics at the University of Toulouse and EHESS, Researcher at Institut D'Economie Industrielle (IDEI) and Industrial Organization Programme Co-Director at CEPR.

<sup>1</sup> This article draws heavily from a report on the economics of railroads made by a team of IDEI researchers led by Paul Seabright. See [http://idei.fr/doc/wp/2003/rapport\\_db\\_1.pdf](http://idei.fr/doc/wp/2003/rapport_db_1.pdf).

<sup>2</sup> Note that in the U.S., railroads own their tracks, they compete for shipments and they rent to competitors the access to their tracks.

infrastructure from operational services.<sup>3</sup> As compared with the traditional model of railway organization where a single firm is in charge of both the fixed infrastructure and the rolling stock management, in vertical disintegration competitors are allowed to offer rail services. The infrastructure remains under the control of a public or private monopolist (which requires some public regulation), but market forces are supposed to be strong enough to generate efficiency in services provision. As in other network industries, the dilemma lies in the organization of the interface between the two separated layers. There exists a strong need for coordination between the infrastructure manager and the users of the infrastructure; this tends to favor integration. And conversely, there is a strong need for competition in services, which argues for disintegration. This probably explains why in most countries where competition has been introduced into rail transport, the solution is “partial disintegration” (Figure 1).

When vertical separation is complete, the main problem is to ensure that the monopolist does not abuse its position: it must be regulated. The partial disintegration case is trickier, since the entity in charge of the infrastructure is simultaneously a provider and a competitor to its challenger. Consequently, it may have some incentives to distort competition in rail services and the public authority faces a complex problem of combined sectoral and competition regulation (Rey et al. 2001).

Depending on the nature and the closeness of the integration between upstream and downstream activities, it may also be more difficult for the authorities to have access to the information required for effective regulation than in the disintegrated case – information about costs, for example. However, a well-

Figure 1

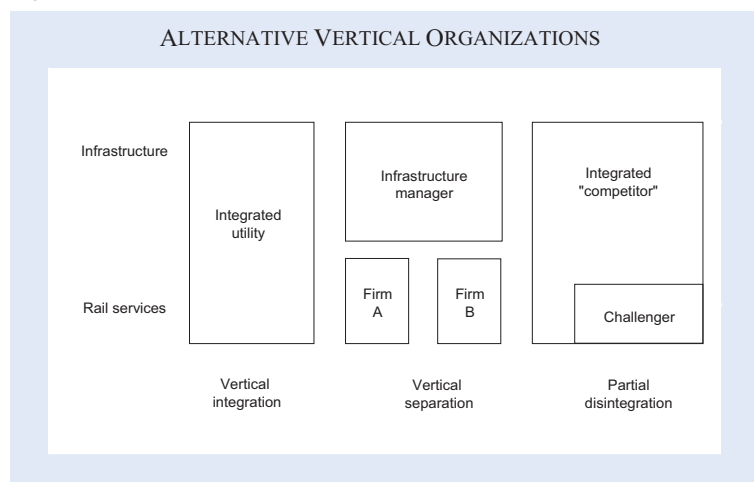


Table 1

Cross elasticity of marginal cost

Freight activity	Infrastructure (t-ratio)	
Intermodal	0.31	(1.33)
Bulk	0.52	(2.62)
General	-0.04	(-0.55)

Source: Ivaldi and McCullough (2001), Table 8.

known advantage of vertical integration is its diminished incentives for double marginalization, so it may be that some kinds of anti-competitive behaviour become less likely under integration even though the authorities’ ability to monitor them is diminished.

What does the empirical evidence show about the effects of vertical disintegration on operating costs? Ivaldi and McCullough (2001) test for cost complementarities in freight transport between infrastructure and operations for US railroads using a translog specification. According to the estimates shown in Table 1, the marginal cost of inter-modal and bulk operation increases with infrastructure output. The negative result for general freight is not statistically significant.

A later article by Ivaldi and McCullough (2002) tests for sub-additivity in the cost function for infrastructure and freight operations. The results indicate that firms running each activity separately have 2.42 percent higher operational costs than a vertically integrated firm. A study by Cantos (2001) undertakes a similar approach to Ivaldi and McCullough (2001) for European services. Using a translog cost function, the author analyzes economies of scope between infrastructure output<sup>4</sup> and transport operations (passenger and freight) for 12 major European railways along the 1973-1990 period. The main finding is that the marginal cost of passenger output is increasing with the level of infrastructure value. The opposite result is obtained for freight operations. As Table 2 shows, the cost

<sup>3</sup> Note that it is not the only type of vertical disintegration. When a trip from A to C necessitates a stop at the intermediary node B, the segment AB can be viewed by the BC operator as a necessary input to provide AC and similarly, the segment BC is essential for the AB operator to provide AC. For this reason, the separation of AC into two products (namely AB and BC) can be considered as vertical disintegration.

<sup>4</sup> The monetary value of all infrastructure facilities (track, buildings, stations, etc.) is employed as a variable for measuring “Infrastructure Output”.

**Table 2**  
**Cross elasticities of marginal operating costs with respect to infrastructure**

Railway undertaking	Passenger marginal cost	Freight marginal cost
	with respect to infrastructure	
BR UK	0.119	-0.052
DB Germany	0.076	-0.143
DSB Denmark	0.132	-0.053
FS Italy	0.905	-0.081
NS Netherlands	0.156	-0.027
NSB Norway	0.133	-0.076
OBB Austria	0.063	-0.116
RENFE Spain	0.082	-0.099
SJ-BV Sweden	0.145	-0.065
SNCB Belgium	0.106	-0.073
SNCF France	0.070	-0.138
VR Finland	0.118	-0.091
Average	0.108	-0.085

Note: All values are statistically significant at 5 percent.

Source: Cantos (2001), Table 5.

anti-complementarity in passenger transport holds for all firms, being more severe for the smaller networks.

Other evidence comes from Mizutani and Shoji (2001), who studied the case of Kobe-Kosoku Railway in Japan. They found that vertically separated firms cost 5.6 percent more than an integrated system.<sup>5</sup> Shires et al. (1999) compared the cost of the Swedish operator after a reform involving vertical separation and found that operating costs had been reduced by 10 percent. However, it is difficult to know to what extent such reductions were due to vertical separation per se rather than to other aspects of the reforms.

In general the evidence, such as it is, is mixed and inconclusive. All studies except one (the Shires et al. study of Sweden) estimate cost complementarities using data from currently integrated firms, which leaves the studies vulnerable to bias due to internal cost-allocation rules, and which means they are unable to take account of what may be the most important effects of vertical disintegration, namely transactions and coordination costs. The UK case suggests these costs may be large, though it is difficult to generalize from a single (and rather unusual) case. Overall, it seems safe to conclude that existing cost studies do show that vertical disintegration of infrastructure from operations could represent a significant cost penalty so that it is wise to account for them if one decides to disintegrate the system, at least to compensate for transaction costs. However,

given that such separation has occurred very rarely to date or more precisely is too recent, the value of such studies in predicting the future consequences of such separations is still limited.

Such studies cannot realistically shed light on one important issue that concerns the role of vertical coordination in influencing the evolution of network structures. In airline networks (unlike rail networks), market entry can create new routes without the need for prior infrastructure investment. To be more precise, provided airport infrastructure exists at the cities at either end of a route, any entrant to the industry can create a direct flight link between two cities where none existed before. However, this cannot happen in railways, where tracks need to be laid before trains can pass. Such entry by airlines has proved of immense importance in shaping the evolution of structures towards hub-and-spoke models in the US, and has begun to be important in allowing new entrants to offer competitive services in the European market. Furthermore, although airport infrastructure can become congested and thereby impose a constraint on network development, the creation of new routes is an important mechanism by which signals of the need for airport infrastructure investment are perceived. In railways, though, network investment will always need to lead rather than lag new route entry by service operators. That implies that the infrastructure operators will need to have much closer coordination (concerning future operation intentions) with service operators than is necessary in the air transport industry. Vertical integration and vertical disintegration with close investment and operational coordination are both feasible options; vertical disintegration with an arms-length relationship between infrastructure and service operators is not.

The literature on “transactions costs” (Williamson 1985) has provided some important insights on the role of vertical integration in industries with high sunk costs of investment (as in railways). For obvious reasons, it is important to ensure that productive investment does not fail to take place because of a lack of coordination of the upstream and downstream parties’ intentions, due to their lack of integration. Williamson’s insight is that such failure may not occur simply because of a breakdown of communications but for a much more fundamental reason, which he terms the “hold-up problem”. Suppose that one party invests prior to the other, and that the investment creates a “specific asset” – one that is worth much less outside the relationship between the two

<sup>5</sup> As reported in Mizutani and Nakamura (2001).



parties. For instance, the asset may be a stretch of railway track that is adapted for high speed trains, which only one operator can run (other operators can run normal trains which do not make full use of the valuable track). Then, as soon as the investment has been irrevocably committed, the HST operator has an incentive to toughen its bargaining position, threatening not to make its own share of the relevant investment. The track operator would have, in effect, to bribe it to invest by lowering the access price to the track, and it might have to lower the price all the way to the price it could charge other, non-HST operators. Naturally, the fear that this might be the outcome would be a disincentive to investing in the track in the first place.

Various possible solutions to the hold-up problem have been proposed, including long-term contracts (which in this example would set the access price at an agreed level even before the track investment had been committed). Long-term contracts can be difficult to write, however, especially when future circumstances may change in unforeseen ways. Instead, vertical integration between track and service operators may resolve the problem by ensuring that neither has the incentive to bargain with the other after the commitment of the investment. Integration does not have to be complete; joint ventures on specific projects by partners that otherwise remain separate are an alternative that may work when the projects are sufficiently distinct. Nevertheless, it is worth bearing in mind that ill-considered vertical disintegration by regulatory fiat may cause difficulties of investment coordination that are not just “communication problems” but go to the heart of negotiation incentives.

To summarize, vertical integration has some disadvantages in a transport network, due to the potentially greater opacity of costs and other operating information that makes effective regulation more difficult, and leads to a risk of anti-competitive discrimination by the network operator against services supplied by a downstream competitor. However, the list of potential advantages of vertical integration is long. It includes some aspects on which empirical evidence is available (notably the extent of vertical economies of scope), and others (notably transactions costs and the risk of hold-up problems) on which evidence is scarce but which may plausibly be extremely important. The overall balance of advantages in vertical network integration is therefore a subject on which further information and research is very much required.

## Competition

Prices are but one of a wide array of business tools that rail firms can use to compete for passengers. In fact competitive strategies concern not just the terms on which a given service is made available to customers but also the choice of the kinds of service to supply, a choice which has a large number of collateral implications for investment, employment policy, and policy towards acquisitions, outsourcing and joint ventures.

Two features of rail travel make consideration of price competition somewhat different from many other industries. The first is that short-run cross-elasticities between transport modes are rather low, suggesting that for rail to compete purely on price against cars or air travel is not likely to yield rapid profits; at any rate, low price strategies would have to be maintained, and seen to be maintained, over a significant period of years before significant traffic could be gained from other transport modes.

The second is that, because of economies of density, price competition that significantly increases traffic can be an extremely profitable strategy for the firm that undertakes it: the true marginal cost of additional traffic lies some way below the average cost. Thus where on-track competition is feasible, or where the characteristics of a given route suggest inter-modal competition may be unusually keen, the incentives to cut prices can be very strong. This has three important implications. First, stable on-track competition may often not be viable: either it is infeasible, or it is feasible and the result is such fierce price competition that unless the competitors have precisely similar cost structures one of them may be forced to withdraw. This may make it quite difficult to support an industry structure with significant amounts of on-track competition, a fact that should be borne in mind in considering regulatory appraisals of the results of introducing competition. We consider this again later.

Secondly, both entrants and incumbents seek for ways to soften the impact of competition by differentiating their products. For instance, non-interchangeability of tickets, non-cooperation over scheduling connecting services, different approaches towards discounting and the targeting of different customer groups, may be tempting strategies for all competitors even if their effect on overall customer welfare is negative. Note that this is quite different from similar strategies used

with predatory intent, in order to drive competitors out of the market. When there is (successful) predation, it is the exiting firm that suffers as well as consumers; when the strategies aim merely at softening competition, the firms benefit and consumers lose.

Thirdly, where inter-modal competition can work (such as on inter-city routes between 200 and 400 km for competition with road and 500 km to 1000 km for competition with air travel) its effect on prices may be important, and may make price regulation unnecessary in circumstances where it might otherwise have been desirable (Antes et al. (2004) and Ivaldi and Vibes (2005)).

For this to be possible, of course, it is necessary that rail services develop characteristics that make inter-modal competition realistic. High-speed trains have done this with some success (though at high cost) in recent years, and it remains to be considered whether and to what extent other kinds of rail service can provide a significant challenge to other modes, notably the car.

Finally, it is worth noting that developments in communications and information technology, notably of course the spread of the internet, are making an important difference to the sophistication of the pricing strategies that firms can adopt. This is particularly true in the realm of price discrimination. Economists distinguish three types of price discrimination.

Both second- and third-degree price discrimination have long featured in rail pricing, notably through season tickets and discounts for the young and the elderly. However, the internet, and information technology more generally, are making sophisticated second-degree price discrimination easier, notably because customers can be shown, quickly and intuitively, the effect of different pricing packages in a way that allows for an informed choice between them. The effect on third-degree price discrimination is more ambiguous. In some respects such discrimination is becoming harder, because customers can shop around, and it is no longer possible to discriminate between customers according to where they are physically located when they make the transaction. However, this is substantially offset by two other considerations. The first, which is particularly relevant to transport, is that when the product is a service that must be consumed at a certain place and/or time, it remains possible to discriminate between customers according to location or time of consump-

tion rather than location of transaction. Airlines have discovered this in a big way through their computer reservation systems, which can charge very different prices for flights of the same length according to the origin or destination, as well as according to the time of travel.<sup>6</sup> The second reason is that firms can now use sophisticated databases of consumer travel behaviour to target special offers to the individual's presumed preferences.

Overall the increased sophistication of price discrimination is likely to make price competition a more tempting prospect, but also to allow competing firms to segment markets in terms of customer types more effectively than has been possible to date.

Then what is the likely outcome of on-track competition? What can be expected to happen to market shares under on-track competition, assuming equal access to any infrastructural facilities that have natural monopoly characteristics? In particular, does equal access imply that there are likely to be reasonably equal outcomes as measured by market share? And conversely, if outcomes are not equal will this imply a failure of equal access? The answer to these questions requires us to look at both the demand and the supply side of the industry.

The reaction of demand depends on the degree of differentiation between the services proposed by the incumbent and the entrants, and on any switching costs that may be incurred when moving from one provider to another. For occasional travelers, there is unlikely to be any switching cost, and competition with newcomers will be tough. For frequent travelers, switching costs may be high and the incumbent will probably keep a large market share independently of cost considerations. An additional argument that implies increased switching costs comes from the network characteristics of passenger rail. Only few people would be able to travel point-to-point with a competitor. The majority of customers would have to change trains, partly using local transport as feed, partly using the incumbent.

Let  $u_i$  denote the utility of a passenger when traveling in a train operated by the incumbent and let  $p_i$  denote the fare. Absent any competitor, the incumbent can charge a price such that  $u_i - p_i \geq \bar{u}$  where  $\bar{u}$  is the net utility from alternative nodes. When there

<sup>6</sup> In actuality, time of transaction (unlike location of transaction) remains a very effective tool of discrimination, because individuals cannot travel freely in time. Tickets booked at the last minute may be very different in price from those booked long in advance.



is an entry, to keep its clients the incumbent has to fix a price such that  $u_i - p_i \geq u_e - p_e$  where  $u_e$  and  $p_e$  are respectively the utility and the price of the service provided by the entrants (and assuming that entry is feasible, which means  $u_e - p_e \geq \bar{u}$ ). The difference  $u_i - u_e$  is the value of the incumbent's advantages: the higher the switching costs, the larger this utility differential. For occasional travelers,  $u_i - u_e$  so that  $p_i \leq p_e$  is necessary to keep these clients, which is feasible only if the incumbent has a cost advantage, that is when  $c_i \leq c_e$ . For frequent travelers, a tariff such that  $p_i \leq u_i - u_e + c_e$  prevents any entry. Even if  $c_i > c_e$ , the switching cost  $u_i - u_e$ , which is positive, allows the incumbent to cut the entrant's price.

Switching costs (natural or strategic) are a strong limitation to competition, as has been well documented in banking, telecommunications, electricity distribution, air transport and pharmaceuticals. As a consequence, incumbent firms will probably keep dominant market shares in subsets of weakly flexible demand, and will lose shares in relatively contestable sub-markets (this may also be a consequence of natural utility advantages of remaining with incumbent suppliers).

The nature of competition is also determined by the nature of the costs faced by the entrant and the incumbent. This affects how low a price each can afford to set in order to attract customers. Other things equal, the lower is the entrant's cost of operations relative to that of the incumbent, the more intense will be the nature of price competition and therefore the higher the likely market share that the entrant can attract.

However, when the incumbent operates a network and the entrant competes only on point-to-point routes, there is an important source of asymmetry induced by network effects. For the entrant, the marginal cost of an additional train full of passengers is the cost of running the train (including administrative costs), plus the access charge for the service. For the incumbent, the true marginal cost consists of the same elements as for the entrant,<sup>7</sup> plus an additional element, i.e., the opportunity cost, which is any additional net cost incurred on those connecting routes to which some of the passengers may subsequently transfer.<sup>8</sup> When transferring passengers in fact yield a profit on the connecting routes, this opportunity cost is negative, and an incumbent's true cost lies below a

<sup>7</sup> These same elements may of course have higher values for the incumbent, for instance if the incumbent faces diseconomies of complexity from running the network.

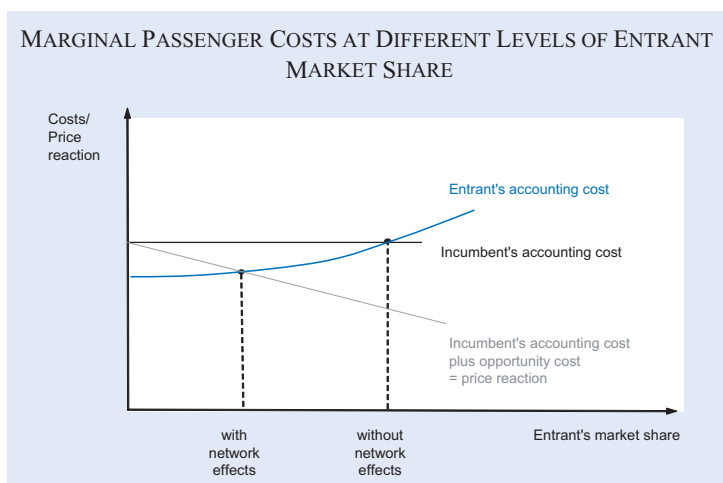
conventionally-measured accounting measure of its costs of providing the service. Three consequences follow from this:

- First, in networks where connecting traffic is a comparatively large fraction of overall traffic (like Germany but unlike France, say), there are fewer cherries for entrants to pick, i.e., there are relatively few connections with a high point-to-point demand.
- Secondly, even when entrants appear to enjoy a cost advantage as normally measured, the opportunity-cost element will mean that the incumbent is a tougher competitor than this advantage would indicate (because it has an incentive to protect its connecting traffic). Where entrants do compete head-to-head with incumbents, their likely market share is usually lower than conventional cost comparisons would lead us to predict.
- Thirdly, it is likely that the Mohring effect (see Small 1992) means that the opportunity cost element becomes more important as the entrant's market share increases, since the reduction in the value of frequency of service to passengers becomes progressively more important as the frequency itself declines.<sup>9</sup> The cancellation of half the services is more costly to passengers if services previously ran every two hours than if they ran every fifteen minutes. This implies that not only will the incumbent's true cost lie below its accounting cost, but it will be significantly more steeply sloped (Figure 2). Therefore even an entrant with a significant initial cost advantage finds that as it eats into the incumbent's market share its cost advantage is progressively eroded. To put it another way, the eventual market share of the entrant is likely to be less sensitive to its initial cost advantage than if opportunity cost considerations did not play a role (in the figure equilibrium market shares are drawn where marginal costs of incumbent and entrant are equal). Without opportunity cost considerations, an entrant with an initial cost advantage could easily reach a large market share, but when opportunity costs matter its market share will be unlikely to become very large. This argument assumes that an entrant's

<sup>8</sup> This is similar in spirit to the opportunity cost calculation that underlies the Efficient Component Pricing Rule for access price regulation (sometimes called the Baumol-Willig rule). The difference here is that we are considering a complementarity between two services on both of which there is competition, rather than a complementarity between shared infrastructure with mandated access and a competitive downstream service.

<sup>9</sup> The Mohring effect is an external effect stemming from the increase of traffic in public transport. This effect leads to a positive externality: When traffic increases, the operator is led to increase the frequency of the services, thus by reducing the waiting time it reduces the total travel time of the users.

Figure 2



services are not perceived by the passenger as contributing to the overall frequency of the service, perhaps because of non-transferability of tickets.

The overall conclusions to be drawn from this line of reasoning are twofold. First, given the unavoidable asymmetry between a network operator and an entrant on point-to-point routes, it may be unlikely that effective competition leads to large market shares for the entrant. This is not, however, to say that we can predict exactly how large such market shares are, since circumstances vary significantly from route to route. Secondly, the share of the incumbent is larger the more polycentric the network and resulting effects are. Thirdly, this asymmetry is not a sign of a market regime failure. Fourthly, the opportunity costs of network traffic are genuine social costs, so that considering the “success” of competition purely in terms of the market shares gained by entrants on point-to-point routes would be seriously short-sighted. If these market shares come at the expense of disruptions in network connections, they may well be symptoms of the failure of competition rather than its success.

Finally, an important point to note is that if the entrant has a significant cost advantage over the incumbent in providing the service, then the incumbent may have an interest in arranging interconnections so as to capture as much of the network traffic as it can. When the costs of interconnection can be avoided, so that passengers can easily switch operators to make through journeys, then the incumbent’s connecting services become complementary to the point-to-point services of the entrant. In these circumstances the incumbent may not only not be damaged but may positively benefit from the cheap fares

provided by the entrant, since these will increase demand for the through journey. This leads naturally to the question how to judge the value of interconnections, which goes beyond the scope of this article.

### Concluding remark

In this article, many questions have not been discussed such as competition for track, the role of regulation, the choice of investments and the structure of the industry that produces equipment for railways. The

economics of railways is a fruitful domain for research in industrial organization.

### References

- Antes, J., G. Friebe, M. Niffka and D. Rompf (2004), *Entry of Low-Cost Airlines in Germany. Some Lessons for the Economics of Railroads and Inter-Modal Competition*, mimeo, Deutsche Bahn and University of Toulouse.
- Cantos, P. (2001), “Vertical Relationships between Railways Operation and the Infrastructure for the European Case”, *Transport Journal*.
- Di Pietrantonio, L. and J. Pelkmans (2004), “The Economics of EU Railway Reform”, *Bruges European Economic Policy Briefings* 8.
- Ivaldi, M. and G. McCullough (2001), “Density and Integration Effects on Class I Freight Railroads”, *Journal of Regulatory Economics* 19(2), 161–182.
- Ivaldi, M. and G. McCullough (2002), “Subadditivity Tests for Network Separation Using a Generalized McFadden Cost Function”, *IDEI Working Paper*.
- Ivaldi, M. and G. McCullough (2005), “Welfare Tradeoffs in US Rail Mergers”, *IDEI Working Paper*.
- Ivaldi, M. and C. Vibes (2005), “Intermodal and Intramodal Competition in the Passenger Transport Market”, *IDEI Working Paper*.
- Mizutani, F. and K. Nakamura (2001), *The Japanese Experience with Railway Restructuring*, Kobe University, mimeo.
- Mizutani, F. and K. Shoji (2001), *Operation-Infrastructure Separation in the Japanese Rail Industry: The Case of Kobe Kosoku Tesudo*, Kobe University, mimeo.
- Nash, C., B. Matthews and J. Shires (2004), *The EU Transport Policy White Paper: An Assessment of Progress*, Institute for Transport Studies, University of Leeds.
- Rey, P., P. Seabright and J. Tirole (2001), “The Activities of a Monopoly Firm in Adjacent Competitive Markets: Economic Consequences and Implications for Competition Policy”, *IDEI Working Paper*.
- Shires, J., J. Preston, C. Borgnolo and M. Ponti (1999), *SORT IT Deliverable VI. Joint Report on the Rail Case Study*, European Commission, Brussels.
- Small, K. A. (1992), *Urban Transportation Economics*, Harwood Academic Publishers, Chur.
- Williamson, O. E. (1985), *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, The Free Press, New York.



## REGULATION AND DEREGULATION IN THE JAPANESE RAIL INDUSTRY\*

FUMITOSHI MIZUTANI\*\*

### Introduction

This paper summarizes regulation in the rail industry in Japan and the current deregulation situation since 1987. One important turning point in Japanese railway regulation policy was the privatization of the Japan National Railway (JNR) and its subdivision into 6 passenger JR companies and 1 nationwide freight JR company (see details in, for example, Mizutani and Nakamura 2004). In 1987, the Railway Business Law was enacted to regulate all kinds of rail organizations, and a decade later, in 1997, the ceiling price regulation was introduced. Although a yardstick competition scheme (or yardstick regulation) had been used for the assessment of 15 large private railways, a more systematic yardstick competition scheme was developed and expanded to apply to 6 JR companies and 10 public subway systems. In 2000, entry into the industry and pricing were largely deregulated. First, as for the entry regulation, the license system for entry was changed to a permission system. Second, whereas the demand-supply balance had been an important criterion in the regulation of entrance into the railway market, that criterion was abandoned. As for price deregulation, it became permissible for a rail operator to change rail fare freely simply by reporting changes to the regulator, as long as the fare change kept the price lower than the ceiling price.

Based on the structure of the regulation scheme mentioned above, I will explain several important points regarding regulation in the Japanese rail in-

dustry. The structure of this paper is as follows. First, I will explain the organization of the Japanese rail industry, focusing on the ownership type and the kind of railways. Second, I will summarize railway regulation, focusing on entry and exit regulation, fare regulation, track fee regulation and so on. Third, I will explain the competition situation, considering the competition for the market and in the market. Last, the yardstick competition scheme, an important characteristic of the Japanese approach to regulation, will be explained.

### Organization of the Japanese rail industry

According to the Ministry of Land, Infrastructure and Transport, there were 199 organizations defined as rail operators, as of 1 October 2003 (Ministry of Land, Infrastructure and Transport 2003). Of these operators, heavy and light rail operators account for 163 operators, and the remaining 36 operators are comprised of monorails, automated guideway transit, cable cars and such.

Passenger rails in Japan are still very important and the share of rail transportation was 27 percent in terms of passenger-kilometers in 2000. On the other hand, the share of freight rail transportation was only 3.8 percent in terms of ton-kilometers in 2000. There are 150 passenger rail operations, but freight rail organizations account for only 13.

Based on Mizutani (1999), we can classify passenger rail operators in four ways: their legal classification, ownership, transport type and main service areas. There are three forms of legal classification in Japan: private corporations, special corporations and public organizations. Private corporations refer to organizations legally considered to be private companies. Private corporations are not always identical to fully privately owned organizations. Therefore, these organizations include companies for which part of the shares are held by the public sector. A public organization is usually a department of the government. These often are the departments of transportation of a local government, for example the Bureau of Transportation in the city of Kobe. A special corporation is

\* Part of this paper was financially supported by research funds from the Ministry of Education, Culture, Sports, Science and Technology and the twenty-first century COE program (Research, Development and Education Center for Advanced Business Systems).  
\*\* Fumitoshi Mizutani is Professor of Public Utility Economics and Management at the Graduate School of Business Administration, Kobe University.

an organization which is settled and regulated by special law (Uekusa 1991). For example, JRs which were established by the privatization of the Japan National Railway are considered special corporations because JRs are still controlled by the special JR law.

Second, as for ownership, there are three categories: private, public and private-public joint ownership. Most Japanese rail operators are privately owned. For example, 15 large private railways such as Tokyu and Seibu in Tokyo and Hankyu and Kintetsu in Osaka, which are widely considered the most efficient railway organizations, are all privately owned railway companies. Public ownership is limited to only 11 operators. Of these 11 publicly owned operators, 9 operators are subway systems such as Tokyo, Osaka and Nagoya. Although three JR companies – JR East, JR Central and JR West – have recently been fully privatized, the other four JRs have not been fully privatized, with most of their shares still being held by the government. Private-public jointly owned organizations comprise what is often called the third sector in Japan. These private-public jointly owned organizations are most often found in small communities.

The most common type of transportation in Japan is urban railways in large metropolitan areas, with the major 15 private railways being classified in this category. However, JRs have both urban and intercity rail services. The most famous bullet intercity rail, the Shinkansen system, is operated by JR companies.

## Regulation

### *General regulations*

The railway industry is highly regulated, just as other public utility industries, such as electricity, gas and water supply, are highly regulated. The Railway Business Law (Tetsudo Jigyoho) has applied to all rail companies since 1 April 1987, when the privatization of JNR was enacted. Before 1987, regulations for JNR differed from those for other railways: JNR was regulated by the Japan National Railway Law (Nihon Kokuyu Tetsudoho) and other railways, such as private railways and public subway systems, were regulated by the Local Railway Law (Chiho Tetsudoho). In addition to the Railway Business Law, there are over 150 laws directed at the industry and enforced by the Ministry of Land, Infrastructure and Transport.

### *Entry and exit regulations*

With the revision of the Railway Business Law in 2000, it was no longer necessary to acquire a rail license before entering the rail market. According to the Railway Business Law (Article 3), if some organization is deemed qualified by the Ministry of Land, Infrastructure and Transport, then the organization can commence rail services. The regulatory change in 2000, therefore, was from a license system to a permission system.

Although the license system has been out of use since 2000, the rail business in Japan is classified into three categories based on the Railway Business Law (Article 2):

- 1) Class 1: enterprises that provide rail passenger and/or freight services while holding their own rail infrastructure;
- 2) Class 2: enterprises that provide rail passenger and/or freight services using rail infrastructure owned by another organization;
- 3) Class 3: enterprises that build rail infrastructure for sale to a class 1 enterprise, or an enterprise which owns infrastructure and rents it to a class 2 enterprise.

While in the European Union rail industry vertical (operation-infrastructure) separation is a common policy, in Japan vertical integration is the norm, with most railway organizations being class 1 enterprises. For example, JR companies such as JR East and JR West, large private railways such as Tokyu and Kintetsu, public subway systems such as Osaka's and Nagoya's, are all class 1 enterprises.

On the other hand, there are few class 2 and class 3 enterprises in Japan. A typical example of a class 2 enterprise is JR Freight. Since privatization, JR Freight has provided rail services by using the rail tracks owned by the six JR passenger companies. As for a class 3 enterprise, Kobe rapid transit railway (Kobe Kosoku) is an example. Although details regarding this organization can be found in Mizutani and Shoji (2004), this company owns rail track and rents it to four private operators, which are class 2 enterprises.

It is worth noting that it is possible for each rail organization to be classified into more than one category. For example, a private railway might have two kinds of rail classes (class 1 and class 2). Although in the case of its Kobe Kosoku line, the Hankyu railway organization is considered a class 2 enterprise be-



cause this line's tracks are owned by the Kobe Kosoku company, for most of its network Hankyu is a class 1 enterprise because it owns most of the rail tracks on which its trains operate.

Criteria for obtaining permission to operate a rail business are described in the Railway Business Law (Article 5). The Ministry of Land, Infrastructure and Transport gives a rail operator permission to provide rail services, if the following four criteria are met:

- 1) The plan is sound from a business point of view;
- 2) The plan is adequate from a safety point of view;
- 3) There are adequate operational plans in addition to the whole plan;
- 4) A potential entrant takes on financial and technological liability.

There are two main new points introduced by the entry regulation of 2000. First, as mentioned before, the license system became the permission system. Although these two systems seem similar, the basic policy has changed. The philosophy of the license system for the rail business was that entry to the market should be prohibited or limited to a very few companies. In contrast, the permission system theoretically grants potential entrance to any organization.

Second, controlling the balance between supply and demand was abolished as an entry criterion in 2000. Before 2000, two criteria related to this item were written in the old Railway Business Law:

- 1) It must be determined that demand for railway service is sufficient;
- 2) There should be no imbalance between supply and demand for rail service when a potential entrant enters the market.

Since the old Railway Business Law was enacted in 1987, there has been some criticism of the criteria, particularly of the supply-demand controlling regulation, which provided no description of the specific conditions necessary to obtain a rail license, such as the minimum demand level and the degree of demand-supply imbalance. The vague and unclear criteria often overprotected incumbent operators. These criteria were abolished because of their effect in deterring competition.

There is no limit to the duration of the permission. Once a rail operator is allowed to operate the rail service, then it is fully responsible for providing rail services, except in cases in which permission is cancelled due to negligence and when the operator exits.

Exit regulations were also lightened after deregulation. The new regulation requires only notification to the Ministry of Land, Infrastructure and Transport one year ahead of terminating the rail services based on the Railway Business Law (Article 28.1). The old regulation required approval from the Ministry of Transport in order for an operator to cease supplying service. This law stipulated that the exit of rail service providers would be allowed if such closure did not damage public interest. This vague description seemed to allow an easy exit for railway service providers, but in fact it was very difficult for railways to go out of business (Saito 1993).

#### *Fare regulations*

The full cost principle is generally applied in the rail industry in Japan. The current regulatory reform in the railway industry has resulted in the introduction of a ceiling price system and a new yardstick competition scheme following the full cost principle which is applied to passenger rail fare (Okabe 2004).

There are five important points regarding fare regulation based on the Railway Business Law (Article 16). First, the ceiling price of rail fare must be approved by the Ministry of Land, Infrastructure and Transport. Second, as mentioned before, because the Japanese railway industry is based upon the full cost principle, the rail fare should cover rail costs including the operator's profits. Generally rail operators are expected not to receive subsidies. Third, as for changes in rail fare, if the rail fare is within the ceiling price, the operator needs only to report the change to the Ministry. Fourth, operators need only to report to the Ministry, when the operators set up new rail fares, such as for express service. Last, if the rail fare increase can be shown to discriminate against some group and/or the fare causes unfair competition with other railways, the Ministry can order the fare to be revised.

The actual application of the ceiling price of the rail fare began in 1997 but was not introduced as a regulation until 2000. Therefore, since 2000, changes in rail fares within the ceiling have become easier because rail operators need not obtain approval from the Ministry.

#### *Rail track fee regulations*

In Japan, there are few rail infrastructure providers (class 3 organizations), so that regulations regarding rail track fees are rather general. For example, the

Rail Business Law (Article 15) rules only that class 1 or class 3 organizations must receive approval of their rail track fees from the Ministry of Land, Infrastructure and Transport. The law does not regulate how much the rail operator pays to the track holding organization, but does require that track fees and conditions of usage be approved by the Ministry.

There are no specific criteria for the assessment of rail track fees. The Enforcement Regulation of the Railway Business Law (Tetsudo Jigyo Shiko Kisoku, Article 30), however, requires the provider to submit documents to be evaluated by the Ministry of Land, Infrastructure and Transport, detailing how the rail track fee is calculated.

There is no single method for establishing rail track fees in Japan. As for JR Freight, the avoidable cost principle is used. But in general, because rail track fees are set up to cover the provider's cost, the Ministry of Land Infrastructure and Transport will consider whether or not the cost is reasonable.

#### *Other regulations*

In Japan private railway operators play a vital role. Therefore, service characteristics and non-rail services are important for private rail operators. First, although service characteristics are regulated by the Railway Commercial Law (Tetsudo Eigyo) and the Regulation for Railway Transport (Tetsudo Unyu Kisoku), the descriptions in these codes are very general. Generally, more concrete rail operation matters such as train schedules can be determined by a railway operator.

Second, Japanese private railways have long been allowed to operate non-rail business as well as rail business. Many private rail companies operate real estate development, retail ventures such as department stores, and other transportation business such as bus and taxi services. However, railway business and non-rail business are strictly separated by Railway Accounting Regulations (Tetsudo Kaikei Kisoku). A railway company is forbidden to allocate rail and non-rail cost at its own discretion, but must follow regulations which describe in detail how to allocate the costs of common facilities and administration. Therefore, it is possible to capture an externality, such as the effect of housing development along rail lines, but an intentional cross-subsidy strategy is avoided, whereby a rail company charges high rail fares and transfers costs from the non-rail service, and vice versa.

#### **Competition for the market and in the market**

Although the license system changed to the permission system for the entry regulation in 2000, there seems to be almost no competition for the rail market because the duration of the effective term of the permission is not stipulated. The system seems based on the concept of traditional monopoly regulation. When the authority gives permission, the authority gives monopolistic rail service to a railway operator while regulating rail fare and service standards to protect rail users from the hazards of a real monopoly. Because the Railway Business Law rules that the transportation committee appointed by the Ministry of Land, Infrastructure and Transport should hear the opinions from the related railway companies and persons when the new railway plan is being considered, there is almost no direct competition for the market.

There are advantages and disadvantages to this system. The advantage of this system is that a rail operator can concentrate on providing better services in the long run, because it protects incumbent rail enterprises from potential entrants as long as incumbents' services are not terribly bad. Although there is no direct competition with potential rail entrants, there is always competition with other transportation modes such as the private automobile. Furthermore, as private railway companies develop areas along rail lines and stations by building housing and operating department stores, they have incentives to provide better services to attract potential rail passengers.

On the other hand, the efficiency of incumbent rail operators may suffer due to a less competitive situation. Rail operators in large metropolitan areas might become particularly complacent because commuter services to large central cities from suburbs are dominated by rail transportation. In order to avoid the inefficiency due to a monopolistic situation, a yardstick competition scheme has been introduced which will be discussed later in this article.

There is almost no competition in the market, which would mean several rail operators competing along the same track. As explained earlier, most rail operators are class 1 operators providing rail services along their own tracks. Of course, there are cases in which a rail company runs trains on a different rail company's tracks. However, most of these are due to the cooperation of two organizations whose best interest is to provide more convenient services for rail users, such as direct train services from suburb-to-suburb through central cities.



### Yardstick competition

One important point about the Japanese rail industry is that the Ministry of Land, Infrastructure and Transport uses a yardstick competition scheme to increase efficiency among existing railway companies. Yardstick competition is seen as competition among rail operators operating in different markets. In this scheme, a regulator sets up several performance measures such as operating cost and evaluates rail operators' performance.

In Japan this scheme has been applied to fare revision in 15 large private railways since 1970s. For example, if a rail operator is inferior to other operators, then as a penalty the Ministry does not approve the fare level desired by the operator. On the other hand, if an operator's performance is better than that of others, then the fare level is approved without revision. Thus any monopolistic behavior due to the licensing system can be counterbalanced to some degree by this scheme.

However, there are two issues to consider with respect to yardstick competition. The first is the question of how effective the yardstick competition scheme is. Although the scheme does not bring about a situation of perfect competition, some kind of competition seems to exist. In fact, Mizutani's (1997) results, based upon Japanese railways' data set, have shown that yardstick competition among large private railways works to some degree. The second issue is related to the number of rail operators involved. The yardstick competition scheme had not been applied to public railways and small private railways until 1997, when the Ministry of Land, Infrastructure and Transport revised the scheme into a more sophisticated tool, making it possible to apply the scheme to three different groups: 15 large private railways, 6 passenger JRs and 10 public subway systems (Okabe 1997, 2004).

In yardstick competition schemes, five measures related to operating cost are used: 1) track costs, 2) catenary costs, 3) rolling stock costs, 4) train operating costs, 5) station operating costs. The standard costs for these five measures are obtained by the following procedure. First, by using data set of each railway, the unit cost for five costs are obtained. The unit cost is defined as the cost divided by appropriate numbers of facilities (e.g. the track cost per track length, the station costs per station). Second, by using these five unit costs as dependent variables, regression analysis is applied. Several variables related to

each cost such as train-kilometers per route-kilometers and numbers of passengers per station are used as explanatory variables. Third, by substituting each rail company's value of explanatory variable into the regression result, the standard unit cost of each rail company is calculated. Next, by multiplying the standard unit cost by each rail company's number of facilities, the standard cost of each category is obtained. Finally, by comparing the actual cost of each rail company with the standard cost of each rail company, the performance of each rail company is evaluated.

The yardstick competition scheme uses an incentive system for rail operators. For the less efficient rail operator, whose actual costs are higher than the standard costs, the reasonable costs for the fare level are the same as the standard costs. Therefore, in the next period, the rail operator is expected to reduce the actual costs to the level of the standard costs. On the other hand, for the more efficient rail operator, whose actual costs are lower than the standard costs, reasonable costs for the fare level are set at half the sum of the actual costs and the standard costs. Therefore, half of the difference between the actual costs and the standard costs is awarded to the efficient rail operator as a reward.

### Concluding remarks

The following are characteristics of the Japanese rail industry:

- 1) Railways are privately owned except for nine public subway systems;
- 2) Operation-infrastructure connection is vertically integrated;
- 3) Entry regulation is based on a permission system;
- 4) Price is calculated by the full cost principle including capital costs;
- 5) Price regulations are based on an approval system for the ceiling price and a reporting system within the ceiling price;
- 6) There is almost no competition for the market;
- 7) There is almost no competition in the market for sharing tracks;
- 8) A yardstick competition scheme is applied for three railway groups (large private, JRs and subway systems).

### References

Ministry of Land, Infrastructure and Transport (2003), *Rail Statistics 2003 (Suji de Miru Tetsudo 2003)*, Unyu Seisaku Kenkyu Kiko, Tokyo, Japan.

Mizutani, F. (1997), "Empirical Analysis of Yardstick Competition in the Japanese Rail Industry", *International Journal of Transport Economics* 24(3), 367–92.

Mizutani, F. (1999), "Changing Trains: Japan", in D. M. van de Velde, ed., *Changing Trains: Railway Reform and the Role of Competition: The Experience of Six Countries*, Ashgate, Aldershot.

Mizutani, F. (2004), "Privately Owned Railway' Cost Function, Organization Size and Ownership", *Journal of Regulatory Economics* 25(3), 297–322.

Mizutani, F. and K. Nakamura (2004), "The Japanese Experience with Railway Restructuring", in T. Ito and A. O. Krueger, eds., *Governance, Regulation and Privatization*, University of Chicago Press, Chicago, 305–36.

Mizutani, F. and K. Shoji (2004), "Operation and Infrastructure Separation: The Case of Kobe Rapid Transit Railway", *Transport Policy* 11(3), 251–63.

Okabe, M. (1997), "New Passenger Railway Fare System: Outline and Features (Atarashii Ryokuyaku Tetsudo Unchin Seido: Gaiyo to Tokushoku)", *Transportation and Economy (Unyu to Keizai)* 57(5), 12–23 (in Japanese).

Okabe, M. (2004), "New Passenger Railway Fares", *Japan Railway and Transport Review* 37, 4–15.

Saito, T. (1993), *Private Rail Industry: Development of Japanese Style Railway Management (Shūtetsu Sangyo: Nihongata Tetsudo Keiei no Tenkai)*, Koyo Shobo, Kyoto (in Japanese).

Uekusa, M. (1991), *Economics of Public Regulation (Koteki Kisei no Keizaigaku)*, Chikuma Syobo, Tokyo (in Japanese).



## RAILWAY PRIVATISATION AND REGULATION IN GREAT BRITAIN

CHRIS BOLT\*

The British<sup>1</sup> railway network was privatised in a process that began over 10 years ago. This paper reviews the process of privatisation and the changes that the railway has undergone and the changing level of state involvement and regulation.

One of the main intentions of privatisation of the network was to create a liberalised structure in which private companies could provide railway services to passengers and freight users, with minimal involvement from the state other than its role in specifying and funding socially desirable services. It was expected that this would lead to increased efficiency and responsiveness to changes in market demands. However, since the initial structure was put in place, there has been further change for the industry, culminating in a Government review published in 2004 and new legislation in 2005.

Throughout these changes, one constant factor has been the Office of Rail Regulation (ORR),<sup>2</sup> the independent economic regulator for the rail industry. ORR was created by the same legislation<sup>3</sup> that allowed British Rail (BR), the publicly-owned company that previously ran the British railway system, to be broken up and sold off. ORR's principal responsibility was and still is the regulation of the monopoly and dominant elements of the railway with particular focus on the main rail network infrastructure manager, Network Rail.<sup>4</sup>

ORR has had to do this against a background of changing Government priorities and industry structures, while maintaining our independence in seeking to promote the public interest as set out in our statutory duties. So ORR acts in a sense like a referee in a game of football, to ensure that all the players play fairly and stick to the rules. However, unlike a game of football, ORR is also able to modify the rules, and we also seek to encourage effective relationships between the different companies operating in the sector to ensure that the needs of rail users are met in a way which offers best value for money.

### Background

The railways in Britain were originally built mostly without state involvement. Much of the nineteenth century infrastructure, particularly structures, is still recognisable today. Private companies built their own networks, often in direct competition with their rivals, on the basis of private Acts of Parliament. The result was a railway that was not centrally planned, but developed where companies thought that there might be a profitable market. Each major company built its own track, operated its own passenger and freight services and designed and built its own rolling stock (often to unique company standards).

The state took control of the railways during the First and Second World Wars, and promoted a reorganisation in 1923 in which most railway operations were merged into four vertically integrated regional companies. The railways were nationalised in 1948, when BR was formed.

BR was never profitable and received annual Government subsidy despite internal efficiency savings and the sale of non-core activities (such as hotels). In the 1960s and 1970s, the size of the network was reduced by about a third. This contributed to reduced costs, but the network still required state funding, and the Government was forced on several occasions to write off debt. Measures designed to focus subsidy on specific services, for example in the 1968 Transport Act, were also ultimately unsuccessful in creating a more commercial environment and improved efficiency.

\*Chris Bolt is Chairman of the Office of Rail Regulation.

<sup>1</sup> This paper covers the situation in Great Britain, excluding Northern Ireland.

<sup>2</sup> Initially known as the Office of the Rail Regulator; it was renamed when a Board structure was introduced in July 2004.

<sup>3</sup> 1993 Railways Act – since amended by the Transport Act 2000, Railways and Transport Safety Act 2003 and Railways Act 2005.

<sup>4</sup> Network Rail acquired the previous network operator, Railtrack, in 2002.

In 1992, the Conservative Government manifesto declared its intention to privatise BR. The aim was to reduce State involvement in the railway, reduce the burden on taxpayers and bring in private sector funding and expertise. The railways were the last of the major privatisations which had started with the privatisation of British Telecom in 1984, and continued with privatisation of the gas, water and electricity networks.

### Privatisation and regulation of the railway

As well as the last, the railways were the most complex and most politically divisive. Unlike the other companies that had been privatised, BR was dependent on Government subsidy to operate and was expected to remain so when in the private sector. The privatisation also involved the biggest restructuring, with one integrated company being replaced by over 100 separate companies, held together through a complex set of contractual relationships, many requiring the approval of ORR and by licences enforced (and where appropriate modified) by ORR.

Three models of privatisation had been considered before deciding on this approach.

The first was to sell off the whole company as a single entity, with no rail competitors (though of course rail is in competition with other modes of transport). This model had been used for the sale of British Gas, but was criticised for moving a monopoly from the public to the private sector, and not introducing real competition. In the case of British Gas, the result was a decade of confrontation with its regulator, leading to numerous references to the Competition Commission,<sup>5</sup> the introduction of domestic competition and the demerger into three successor companies covering gas exploration, transmission and distribution, and gas supply.

The second model was to break BR into several regional companies, in some ways replicating the situation before the railway was nationalised in the 1940s. This is essentially the structure that still exists in the privatised water industry in England and Wales. This model, which was also used for the privatisation of Japan Railways, was rejected for creating regional monopolies and not introducing competition.

The third option, the one that was adopted, was the “Track Authority” model. BR was vertically separated, with a new private company, “Railtrack”, owning the infrastructure and charging train operators to have access to the network. Although there was initially a view that Railtrack might be retained as a public sector company, it was floated on the stock market in 1996. The track and infrastructure maintenance activities of BR were divided into several companies and also sold off, mostly to construction companies, with contracts to supply services to Railtrack. In parallel, franchises to run passenger services for periods of between 5 and 15 years were also tendered by a public sector body, the Office of Passenger Rail Franchising (OPRAF). Some of these franchises were awarded to management teams, but many were awarded to bus companies.

Railtrack was a monopoly and, as with other privatised infrastructure companies, the privatisation legislation created an independent economic regulator to ensure that it did not abuse its monopoly position. ORR was given the role of regulating the company through enforcing conditions set out in the licence the company was required to have to operate the network.

As part of its regulation of Railtrack, ORR also needed to establish the charging framework that determined the track access fees that the company could set. The charging framework would allow Railtrack to recoup its efficient costs as well as providing a return on its capital. The framework of track access charges is generally fixed for a five-year period, after which it was to be reviewed by the regulator. Following the initial establishment of the charging framework in 1995, a periodic review took place in 2000, with a special interim review taking effect in 2004. The next periodic review of charges will take effect in 2009.

ORR also has a role in ensuring that the network is being operated effectively and fairly in approving all track access contracts. In addition, it has concurrent competition powers for railway services with the Office of Fair Trading (the national competition authority).

At the time of privatisation, there was considerable debate about the appropriate extent of competition between train operators. To be eligible for subsidy, passenger services have to be operated under a franchise agreement. There were initially 25 franchises, with franchises being awarded to operators that of-

<sup>5</sup> Previously the Monopolies and Mergers Commission.

ferred the best value, generally in terms of the lowest level of subsidy (or in a few cases the highest premium). Except where franchises overlapped (with, for example, services from London to Birmingham being provided by three different operators), there were initially restrictions on the ability of so-called “open access” operators to seek access rights to run competing services. To date, there has been only one open access operation of any size, which runs services without subsidy from Hull to London (a route not served well by the relevant franchise).

In contrast, there has been competition in the provision of rail freight services from the start. The freight division of BR was sold off as six separate companies, with the individual companies free to negotiate access with Railtrack. Although in the event five of these companies were acquired by the same owner, other operators have since entered the freight market and have been successful in winning business from other freight operators, as well as bringing new business to rail.

Although Railtrack and all train operators require a licence to provide services, other companies spun out of BR generally do not. These included the companies providing maintenance and renewal services to Railtrack, and the Rolling Stock Companies (ROSCOs), created to own all rolling stock and finance new rolling stock, and lease it to train operators. Three ROSCOs were created at privatisation, and their activities are “regulated” only through general competition law.

### **Restructuring of the industry**

The first years of the railways in the private sector were generally positive. Passenger numbers grew, as did the amount of freight carried. Additional services were introduced on the network. New passenger freight and freight rolling stock were also introduced. Passenger-kilometres have increased by 35 percent from 1994–2004 and freight tonne-kilometres increased by 40 percent over the same period.

Railtrack’s initial performance as a private sector company also appeared to be positive. However, a series of events began to undermine confidence in the company. The company embarked on a major development of the West Coast main line, the key strategic route to Scotland from London via Birmingham, Liverpool and Manchester. The development of that

route was intended to renew and upgrade the infrastructure, resulting in faster journey times, improved reliability and increased capacity. The initial project was very ambitious, and the target improvements were delayed and were subject to significant cost overruns.

Although passenger numbers grew, the franchises were also not without problems. Some of the earliest franchise bidders found that they were unrealistic in their projections. Several franchise operators sought to renegotiate their franchise agreements, and some were allowed additional subsidy.

A major feature of the British rail network in the ten years since privatisation has been the increase in costs. The rise in costs has been attributed to several factors, including higher safety standards, new legislative requirements for example in respect of accessibility of trains and stations, growing risk aversion, and poor cost control and asset knowledge on Railtrack’s part. This was in some part due to the way in which maintenance was contracted out, which resulted in a loss of control by Railtrack of key information on its assets. ORR recognised this last failing and strengthened the company’s licence; however, Network Rail decided in 2004 that maintenance should be brought back in-house, and this transition has now been completed.

Public confidence was eroded in the railways in general and Railtrack in particular by a series of fatal accidents between 1997 and 2000. Following a fatal accident in 2000 at Hatfield, north of London, a large number of speed restrictions were placed on the network. Performance suffered dramatically, costs escalated and Railtrack’s share price fell. This culminated in 2001 when Railtrack was placed into Railway Administration and was ultimately acquired by Network Rail, a company limited by guarantee, owned by about 115 “members” rather than by shareholders.

Before this, the change of Government in 1997 had led to a review of the structure and regulation of the railway industry. The new Government concluded that the public sector structure that it inherited was not fit for the purpose. It created an agency to plan the long-term future of the railway, the “Strategic Rail Authority” (SRA), which took on the franchise award and monitoring functions of OPRAF. The creation of the SRA was an acknowledgment by the Government of the need for a longer term strategy, while retaining the structure of private sector provision put in place at the time of privatisation.



In 2004 the Government announced a further Rail Review, prompted by concerns about the deterioration in performance and escalation in costs since the Hatfield accident and reflecting the other changes in the industry, including the replacement of Railtrack by Network Rail. In the conclusions to the review, published in July 2004, the Government decided to change again the responsibility for rail strategy, bringing this back within central government. As a consequence, the SRA was wound up and its franchising functions were also brought inside the Department for Transport. Network Rail was given additional responsibility for the overall operation and performance of the railways, as well as responsibility for developing Route Utilisation Strategies – medium term plans for each part of the network reflecting increasing demands and the steps needed to meet that demand while maintaining and improving performance. ORR modified Network Rail's licence in 2005 to ensure that it carried out these new responsibilities effectively.

The review confirmed the Government view that independent economic regulation of the industry was essential – a view shared by the majority of the rail industry. The review also concluded that safety regulation of the railways should transfer to ORR from the Health and Safety Executive. Her Majesty's Rail Inspectorate will therefore transfer to ORR in the early part of 2006, so that ORR becomes a combined safety and economic regulator.

The aim of these changes is essentially two-fold, which is recognised in the description of the railway as a “public and private partnership”, involving “public specification and private delivery”. The new arrangements reinforce the responsibility of Network Rail and train operators to work together to deliver improved performance and efficiency in a way which meets the requirements of rail users. But they also place a clear responsibility on Ministers – a responsibility now reflected in the Railways Act 2005 – to set out a “High Level Output Specification” of the rail services it wishes to fund, and the funding available, as an input to ORR's periodic review of Network Rail's outputs and allowed revenue.

### **The role of ORR**

The review therefore confirmed the continuation of a “triangular relationship” between the Government, Network Rail and the franchised train operators with

ORR playing a key role in ensuring that these relationships operate effectively and fairly.

The relationship between Network Rail and the train operators is set out in access contracts. ORR reviews and approves these contracts as well as any changes to the Network Code, the industry wide agreement that is part of each access contract that ensures that the interests of all parties are taken into account in the daily operation of the railway. The access contracts also set out the details of the performance regime. The performance regime is the part of the access contract that penalises poor performance with financial penalties. If a service is delayed due to a failure of the infrastructure, then the infrastructure manager compensates the operators that are affected. The same is also true of delays caused by other operators.

The relationship between the Government and Network Rail is another key feature of the railways and is described in the review conclusions as a “binding arrangement” to ensure that Government gets value for the public money it is putting into the railway. But this is not a contractual relationship. As explained above, the Government determines what level of support it will give to the railway and what its priorities for service delivery are. The outputs the Government wants are set at a high level and will in future take account of the Route Utilisation Strategies developed by Network Rail. Detailed delivery plans are then established for Network Rail through its regulatory framework monitored and enforced by ORR and by the franchised passenger operators through franchise agreements.

The third leg, between the government and the franchised passenger train operators, is the one where ORR has the least direct involvement, but still plays an important role. This relationship is determined by the franchise agreement which sets out the service levels required from each operator and the level of subsidy (or return to the government) that each franchise require. Each train operator also has a licence granted by ORR and subject to a set of conditions. ORR also ensures that the needs of other users of the railway, especially freight, are not ignored. ORR achieves this by approving all access agreements to the network and determining the framework for access charges. As part of this role, ORR has reviewed the format of access agreements and also encouraged the industry to review the multi-lateral arrangements in the Network Code to ensure that responsibilities of train operators and of



Network Rail are clear and that all companies involved in providing services to rail users can work together to deliver improved performance and efficiency. These changes have continued following the rail review, for example with a move towards integrating Network Rail and train operator control rooms, to allow more effective real-time management of train services.

Taken overall, the changes in the structure and regulation of the rail industry in Britain anticipated much of the EU legislation that aims to reform and revitalise the railways. For example, rail privatisation introduced the vertical separation of the railway industry (though EU provisions do not require privatisation), the encouragement of new entrants into the market, the licensing of railway undertakings and the creation of an independent regulator to act as appeal body. These regulatory bodies have now been created in each member state that has a railway and are cooperating with each other to ensure that international, as well as domestic, rail services are operated in an environment that is fair and open. In Britain the new European legislation extends our powers to previously unregulated facilities, such as the Channel Tunnel Rail Link.

### **Conclusion**

The privatisation of BR transferred the ownership of the railways from the public to the private sector. At the same time the government created a regulatory structure that would ensure that the railways fulfilled the various public obligations that were required of it. The initial structure has been amended several times since privatisation to ensure that it is fit for purpose. In particular, it is important that different companies operating in a network industry work together to deliver the performance and efficiency that customers rightly demand. This requires an appropriate set of contracts and licences, and effective partnership, as now exists in Great Britain.

Our vision is for the mainline railway industry in Britain to be one in which Network Rail, train and freight operating companies in partnership with public sector funders, and the railway supply industry, work together to meet the current and future needs and reasonable expectations of passengers and users by providing safe, high performance, well maintained and efficient railway services that offer value for money for passengers, other users and funding organisations. That vision underpins our approach as the independent economic – and in future also the safety – regulator for the railway.

## RAILWAY (DE-)REGULATION IN GERMANY

GÜNTER KNEIPS\*

### Railway (de-)regulation in the context of EU transport market liberalisation

Since the European Court of Justice ruled against the Council of Transport Ministers in 1985 for failing to ensure freedom to provide services in the sphere of international transport, the paradigm shift towards full competition on the European transport markets has become irreversible. The European Union has played a leading role in this process, and the benefits of free entry to transport markets throughout Europe are now largely unchallenged.

The transportation of persons or goods on roads, railways, waterways, and in the air, seems at first glance very heterogeneous. Nevertheless, the different transportation modes share as a common denominator the existence of an infrastructure of routes, of traffic control systems as well as vehicles to provide transportation services. For example, railroads are technical systems which can be divided into the following related parts:

- Tracks and stations (construction and maintenance);
- Train traffic control systems (scheduling and operating);
- Train services (transportation of goods and passengers).

There are obviously strong complementarities between the different parts of railroad systems. Train services can only be provided if access to tracks and stations is guaranteed and the operation of trains is coordinated, including ex ante scheduling as well as real time train control. These synergies have created the fable of vertical integration as the adequate or-

ganisational form of railroad systems. Lessons from the history of the nineteenth century that it might be more effective to organise railroad systems in a rather disaggregated way have been widely ignored.<sup>1</sup> It is by now well known that third-party access to railways is indeed technically feasible. Indeed, the process of regulatory reform during the last decade would otherwise have been pointless.

The disaster of reduced traffic and increased deficits of European railroad companies has led to a challenge of the vertical integration approach. The principles of non-discriminatory access charges to railway infrastructures were already laid down in the Council Directive 91/440/EEC<sup>2</sup> of July 1991 as the precondition for the competitive supply of railway services on the same track. Free entry of service companies should improve the quality and variety of train services as well as provide incentives for a more cost-efficient production of train services. Vertical integration is no longer considered to be the adequate organisational form of railway systems. Instead, EU policy has been to separate the supply of train services from the provision of infrastructure, separation of accounts being compulsory and organisational or institutional separation being optional. In the Council Directive 95/19/EC of 19 June 1995 the basic principles of infrastructure allocation were established on the Community level.<sup>3</sup> These principles do not allow discrimination between national and international services, discrimination between different users of railway infrastructure and excessively high access charges. The design of the non-discriminatory allocation of track capacities, however, remained within the competence of the member countries. In particular, market power regulation has not been prescribed by the EU Directives.

Efficient competition on European rail transport markets is conditional upon the existence of non-dis-



\*Günter Knieps is Professor of Economics at the University of Freiburg and Director of the Institute of Transport Economics and Regional Policy.

<sup>1</sup> In the specific German case, generations of transport economists had regarded the nationalization of the Prussian railways from 1879 onwards as the logical and cogent solution for railway systems as such, disregarding the fact that even then alternatives existed (Fremdling and Knieps 1993, 129).

<sup>2</sup> The Council Directive 91/440/EEC of 29 July 1991 on the development of the Community's railways, OJ L 237, 24.08.1991.

<sup>3</sup> The Council Directive 95/19/EC of 19 June 1995 on the allocation of railway infrastructure capacity and the charging of infrastructure fees, OJ L143/75, 27.06.1995.

criminatory access to rail infrastructure for all active and potential train service providers. In addition, however, efforts must also be made to ensure that scant infrastructure capacities are allocated efficiently and total costs of rail infrastructure are covered.

### **The German railroad reform**

On 1 January 1994, the railway reform legal package was enacted. The transition from a public enterprise to a firm under private law in the form of a joint stock company can only be called formal privatisation (rather than a real privatisation by sale of publicly owned assets), because the state is still the sole owner of the Deutsche Bahn AG (Group). Separate branches for infrastructure (DB Netz AG), commodity transportation (DB Transport und Logistik), passenger long-distance transportation and passenger local transportation were founded. Financial reasons also played a non-negligible role for the privatisation initiative. The Deutsche Bundesbahn and the Deutsche Reichsbahn, its counterpart in East Germany, suffered from large amounts of debts. The first step of the privatisation thus consisted of the relief of the liquidation of debts and the endowment with new capital.

There has been an intense controversy over the issue of separating railway infrastructure from railway services and not only formally privatising the service companies of the Deutsche Bahn AG (Group) (Knieps 1996, 44; Wissenschaftlicher Beirat beim Bundesministerium für Verkehr 1997, 632). Such a real separation, however, has not taken place so far. Until now, German railways have only been formally privatised. Deutsche Bahn AG (Group) is the only shareholder of the subsidiaries mentioned here. The Federal Republic of Germany has so far remained the only shareholder of Deutsche Bahn AG (Group). Since private capital can only be raised if risk-equivalent interest rates can be expected, privatisation shifted public attention to the cost-covering possibilities of access charges to the rail infrastructure.

A major goal of the German railroad reform has been entry deregulation of train services in the context of the liberalisation of European transportation markets. Accounting separation between service level and infrastructure level was considered a necessary precondition to guarantee non-discriminatory access to the tracks for all providers of train services. The DB Netz AG is obliged to provide access to the service providers' tracks on a non-discriminatory basis. Access charges have to be paid by all users of the

infrastructure. To cover the gap between revenues from track access charges and total costs, the State contributes to the financing of the infrastructure.

### **First phase of (de-)regulation: Negotiated third party access**

If the track owner no longer supplies all transportation services himself, it is vital to distinguish between the service tariff the customers have to pay to the transportation firm and the access charge the transportation firm has to pay to the track owner. According to the well-established subsidiary principle, in Europe only the basic principles of infrastructure allocation were established on the community level. These principles do not allow discrimination between national and international services, discrimination between different users of railway infrastructure or excessively high access charges. The detailed design of non-discriminatory allocation of track capacities, however, remained within the competence of the member countries.

The first phase of German railway (de-)regulation has been characterised by the requirement of non-discriminatory third party access without ex ante sector-specific regulation. According to the General Railway Act (AEG), §14, all railway companies located in Germany have the right of non-discriminatory access to railway infrastructures, irrespective of the kind of rail transport they offer. The design of three subsequent access charge systems as well as the allocation of track capacities was left within the competency of DB Netz AG. The newly founded Federal Railway Administration was in particular responsible for technical regulations, whereas the competency for issues of access discrimination was increasingly handed over to the Federal Cartel Office. The basic concept was based on negotiations between applicants and the DB Netz AG in its function as infrastructure manager. There was no ex ante regulation of access charges.

The access charges of the DB Netz AG were factually unregulated. The Federal Railway Administration was only responsible for settling conflicts between the DB Netz AG and third parties arising in the context of access conditions and access charges. Earlier criticism of the access charges policy of the DB Netz AG already indicated that the overall level of the access charges would be too high, in particular due to the overload of employees at the Deutsche Bahn AG (Group) (Aberle and Brenner 1994, 707 f.).

DB Netz AG issued its first access pricing system on 1 July 1994, consisting of separate catalogues of prices and conditions for access to its tracks for passenger transport and for freight transport. Its major characteristics were quantity rebates, based on the total amount of train kilometres undertaken on the track network of DB Netz AG. Its successor, the second access pricing system (TPS 1998) was issued by DB Netz AG in June 1998. This revised rail track tariff system featured a two-tier level of charges. After obtaining an “InfraCard”, the track user was charged a lower variable price or, on the other hand, without using this card, he was charged higher rates according to the actual services made use of. Within each demand group, rail track users were treated on equal terms. In the case of capacity constraints arising from the sheer volume of rail track usage, customers using “InfraCards” or the “VarioPreis” (variable charges system) were treated equally. The third access pricing system (TPS 2001), issued by DB Netz AG in 2001, was characterised by a linear tariff without volume discounts or optional “InfraCard”. Instead, elements of product differentiation in the form of different categories of track capacities are offered.

So far, revisions of the access charge systems of the DB Netz AG only seem to occur in reaction to public debate. In particular, the argument that quantity discounts or non-linear tariffs would unilaterally favour the position of Deutsche Bahn AG (Group) as the dominant supplier of rail transportation services and conditions of equal access to the tracks would therefore be disturbed has led to the introduction of linear access charges, which are obviously inadequate to attract more traffic to the railway systems (Knieps 1998, 466 f.).

Active competition on the German railroad market is focused on commodity transportation within Germany as well as on local passenger transportation. Entry into cross-border transportation can rarely be observed; cabotage on foreign networks within other EU countries is almost nonexistent. Competitive bidding for subsidies for local passenger transportation takes place only to a limited extent (Aberle and Eisenkopf 2002, 68).

Deutsche Bahn AG (Group) is the largest provider of rail services in Germany. Based on mileage, by the end of 2003 its market share was 91 percent for local and regional passenger services, more than 99 percent for commercial long distance and interregional

passenger services, and 94 percent for freight services (NEA 2005, 13).<sup>4</sup>

Since the reform of the railway sector there has been almost no entry of commercial long-distance and interregional passenger operators in the German rail market. Few examples are: two lines with low frequency run by Connex (“InterConnex”) and one international night train run by GVG. The market share of the competitors of Deutsche Bahn AG (Group) in non-commercial passenger services was less than 10 percent in 2003, including direct awards to federal state owned railway companies.

More entry can be observed in the German rail freight market. Although Railion (former DB Cargo) is still the dominant operator for freight (>91 percent in 2003), there are other private operators emerging in specific freight markets. The four largest long-distance providers are Railion (Stinnes-Logistics/Deutsche Bahn Group), TX Logistics, HGK/SBB Cargo (Co-operation), and Rail4Chem. In passenger operations the four largest are DB Regio AG, DB Reise & Touristik AG, Arriva Deutschland GmbH, and Connex Regiobahn GmbH (Connex Group) (NEA 2005, 13, 37).

### **The second phase of (de-)regulation: Introduction of market power regulation**

The shift again towards market power regulation of rail access, which was initiated by the EU Directive 2001/14 of the railroad infrastructure package<sup>5</sup> of February 2001, introduces several regulatory obligations for the provider of track access and requires a regulatory body to be set up in each member state. In Germany, a new regulatory authority, the Federal Network Agency (Bundesnetzagentur) has been established. It is responsible for sector-specific regulation for the telecommunications and postal sector, the electricity and gas sector, and the railway sector. According to article 3, detailed statements of the infrastructure provider are required, including details of the charging system and the principles and criteria

<sup>4</sup> Similar figures are also presented in Lindemann (2004, 122).

<sup>5</sup> The Rail Infrastructure Package contains 3 Directives: Directive 2001/12/EC of the European Parliament and of the Council of 26 February 2001 amending Council Directive 91/440/EEC on the development of the Community's railways, OJ L75/1, 15 March 2001; Directive 2001/13/EC of the European Parliament and of the Council of 26 February 2001 amending Council Directive 95/18/EC on the licensing of railway undertakings, OJ L 75/26, 15 March 2001; Directive 2001/14/EC of the European Parliament and of the Council of 26 February 2001 on the allocation of railway infrastructure capacity and the levying of charges for the use of railway infrastructure and safety certification, OJ L 75/29, 15 March 2001.

for capacity allocation (Annex I). Train service companies have the right to appeal to the regulatory agency against decisions of the track provider (Art. 30). The railroad package neither prescribes tariff structures nor enforces price regulations. It leaves a large scope of discretionary power to the regulatory agencies of the member countries.

In the meantime a new Infrastructure Utilisation Regulation has been passed in Germany.<sup>6</sup> Based on the new EU Directives, a set of detailed requirements has been specified in order to improve the transparency of the principles and criteria for the allocation of track capacities as well as the principles of access tariffs. Negotiations concerning the level of infrastructure charges will in the future only be permitted if they are carried out under the supervision of this regulatory body.

During the first phase of (de-)regulation the debate on access charging seemed to neglect the remaining regulatory problems. An essential characteristic with respect to the supply of train services is its network structure. Incentives may exist for train companies for bundling traffic either on a given line (economies of scale) or in serving several lines jointly (economies of scope). Nevertheless, if in a particularly sparsely populated area there is a lack of competition between active firms in the market, this may be replaced by efficient potential competition. The pressure of potential competition is sufficient to create incentives for the active supplier of train services to produce more efficiently. Thus the actual number of active competitors is of negligible relevance, as long as potential entrants can play the role of disciplining the active providers. Therefore, the condition for the functioning of potential competition for disciplining firms already in the market is that the incumbent firms do not have asymmetric cost-advantages compared to potential entrants. Whereas active and potential competition of transportation firms acting on the track initiates a trend towards cost-oriented transportation tariffs, railway tracks themselves must be regarded as monopolistic bottlenecks. The theory of monopolistic bottlenecks is central to the disaggregated regulatory approach in terms of locating network-specific market power in connection with the efforts to determine the minimum basis for regulation (Knieps 1997a, 327–31; Knieps 1997b, 362–68).

<sup>6</sup> Verordnung zum Erlass und zur Änderung eisenbahnrechtlicher Vorschriften vom 3. Juni 2005, Bundesgesetzblatt Jahrgang 2005, Teil I Nr. 32, ausgegeben zu Bonn am 13. Juni 2005, 1566–77.

The conditions for a monopolistic bottleneck facility are fulfilled

- 1) If the facility is necessary for reaching consumers, that is, if no second or third such facility exists, i.e. if there is no *active* substitute available. This is the case if there is, due to economies of scale and economies of scope, a natural monopoly situation, so that one supplier can provide this facility at lower cost than several suppliers; and
- 2) If at the same time the facility cannot be duplicated in an economically feasible way, that is, if no *potential* substitute is available. This is the case if the costs of the facility are irreversible.

The special focus of regulatory activity should be on the design of a symmetrical regulation of the access to monopolistic bottlenecks, combined with a regulation of access charges. Therefore, if a potential competitor plans an entry with a parallel track, the incumbent railway owner could reasonably threaten to reduce his tariffs to the short-run variable costs. Once a railway network is completed, one cannot expect further entries with additional tracks. The decision-relevant costs of entry include the costs of tracks, which cannot be covered by tariffs based on short-run variable costs. In contrast to the supplier of rail-services, the track owner in question has therefore obtained market power. Since competition among lines is lacking, unregulated access charges create the danger of the track owner exploiting his monopoly power.

The shift towards sector-specific ex ante regulation of access to the track seems necessary in order to discipline the impact of market power on the bargaining for access conditions. In contrast to competitive networks, the market power involved in network infrastructures characterised as monopolistic bottlenecks fundamentally disturbs such bargaining processes. One extreme alternative could be (vertical) foreclosure of competitors on a complementary service market. A tying of this sort can be used as a method of price discrimination, enabling a monopolist to earn higher profits. Another way of abusing market power within the bargaining process on access conditions is to provide insufficient network access quality or demand excessive access charges.

An adequate starting point for regulatory intervention when market power is involved in access processes seems to be the “essential facility” doctrine. Well known and often applied in US antitrust law, the essential facility doctrine gains increasing impor-



tance also in European competition law. The focus is on access to monopolistic bottlenecks on equal terms for all competitors. It is through the application of the 1890 Sherman Act that the essential facility doctrine has developed in the US.

Liability under the essential facilities doctrine is based on the following criteria:

- 1) Control of an essential facility by a monopolist (endowing monopoly power);
- 2) A competitor's inability, in practical or reasonable terms, to duplicate the facility;
- 3) The denial of the use of the facility to a competitor; and
- 4) The feasibility of providing the facility.

In the context of the disaggregated regulatory approach the essential facilities doctrine is no longer applied case by case – as is common in US antitrust law – but to an entire class of cases, namely, monopolistic bottleneck facilities. The design of non-discriminatory conditions of access to essential facilities must be specified in the context of the disaggregated regulatory approach.

Whereas ex ante regulation of access to railroad tracks seems necessary, this should, however, not lead to over-regulation. It is important to differentiate between the price level, which has to be regulated, and the pricing structure, which must remain unregulated. Regulators should neither be allowed to prescribe pricing roles that focus on tariff structures within monopolistic bottlenecks nor to forbid per se the implementation of non-linear tariffs. Price cap regulation in the monopolistic bottleneck areas and accounting separation are necessary for disciplining the remaining market power and ensuring non-discriminatory access. Detailed input regulation contradicts the spirit of a price cap regulation. Not only in competitive subparts of networks, but also in the monopolistic bottleneck areas pricing structures should be flexible and the result of endogenous market processes. The welfare-increasing effects of price differentiation should not be impeded by asymmetrical regulatory intervention.

## References

- Aberle, G. and A. Brenner (1994), "Trassenpreissystem der Deutschen Bahn AG – eine erste kartellrechtliche und ökonomische Beurteilung", *Internationales Verkehrswesen* 46, 707–12.
- Aberle, G. and A. Eisenkopf (2002), *Schienenweg und Netzzugang – Regulierungsprobleme bei der Öffnung des Schienennetzes und wettbewerbspolitische Empfehlungen zur Gestaltung des Netzzugangs*, Deutscher Verkehrs-Verlag, Hamburg.
- Fremdling, R. and G. Knieps (1993), "Competition, Regulation and Nationalization: The Prussian Railway System in the Nineteenth Century", *Scandinavian Economic History Review* XLI(2), 129–54.
- Knieps, G. (1996), *Wettbewerb in Netzen – Reformpotentiale in den Sektoren Eisenbahn und Luftverkehr*, Tübingen.
- Knieps, G. (1997a), "Phasing out Sector-Specific Regulation in Competitive Telecommunications", *Kyklos* 50(3), 325–39.
- Knieps, G. (1997b), "The Concept of Open Network Provision in Large Technical Systems", *EURAS Yearbook of Standardization* 1, 357–69.
- Knieps, G. (1998), "Das neue Trassenpreissystem: Volkswirtschaftliche Vorteile eines zweistufigen Systems", *Internationales Verkehrswesen* 50(10), 466–70.
- Lindemann, K. (2004), "Auf den Rahmen kommt es an", in *Wettbewerb im Schienenverkehr – Formen, Stärken, Schwachstellen und Grenzen*, Schriftenreihe der Deutschen Verkehrswissenschaftlichen Gesellschaft e. V., Reihe B, B 274, Berlin, 121–27.
- NEA Transport Research and Training (2005), *Erail Monograph Germany*, version 6, Rijswijk, The Netherlands.
- Wissenschaftlicher Beirat beim Bundesministerium für Verkehr (1997), "Bahnstrukturreform in Deutschland – Empfehlungen zur weiteren Entwicklung", Stellungnahme vom November 1997, *Internationales Verkehrswesen* 49, 626–33.





## RAILWAY (DE-)REGULATION IN EU MEMBER STATES AND THE FUTURE OF EUROPEAN RAIL

JAN SCHERP\*

### Introduction

For around 30 years railway transport has been in relative decline in comparison to other modes. This is particularly well illustrated by the development of rail freight. The European freight transport sector enjoyed vigorous growth in the last decade. The traffic performance rose by roughly 2.5 percent per annum outstripping growth in GDP by approximately 0.5 percentage points and thereby underlining the very high freight transport intensity of economic growth. However, the rail freight sector was unable to take part in this strong growth. Its share in the five transport modes (in the European Union of 15 member states before the last enlargement) fell from 21 percent in 1970 to about 8 percent in 2002. For passenger transport the trend was similar: rail's modal share dropped from 10.5 percent to around 6 percent over the same period.

Traditionally integrated, public monopoly companies provided rail transport services in Europe after the Second World War. Within this framework, the railway sector was unable to respond adequately to challenges such as the globalisation of transport logistics, the shift away from heavy industry towards a service and retail economy, in addition to the increase in car ownership and road building. Rail proved to be unable to take up the challenge of competing with other transport modes that had opened up to competition and became more flexible in adapting to a changing demand pattern. The need to reform the railway sector had grown as a result of the increasing financial

problems of railway firms. Rail suffered additionally through the particularly strong public influence on railway management. It had to deal with many conflicting public interests (transport policy, labour policy and regional development) with few incentives for managers to meet market requirements.

At the end of the 1980s and in the 1990s some member states of the European Union started to restructure the railway sector and to reform the regulatory framework in order to open up rail markets. At the European Community level, the White Paper in 1996 on rail transport (European Commission 1996) laid down the strategic principles aimed at revitalising the railway sector in order to increase its competitiveness and attractiveness with customers. Ideally, the railway sector must be able to participate in transport growth by allowing railway firms to act as commercial entities at a European level. The Community Transport White Paper of 2001 (European Commission 2001) confirmed this approach and defined a political target of maintaining the 1998 rail modal share by the year 2010. In order to reach this target the Community rail policy aims at:

- Ensuring non-discriminatory market access and transparent market structures,
- Providing incentives for an efficient infrastructure use,
- Contributing to a sustainable financial restructuring of railway undertakings and infrastructure managers,
- Triggering a positive rail market development.

After the White Paper of 1996 the European Commission has launched, from the late 1990s onwards, legislative initiatives to integrate the rail service markets and aimed at developing a common European railway area. Key elements of this initiative are a gradual market opening of rail freight and international rail passenger services, promoting the interoperability of the various national systems through the gradual implementation of common technical specifications of interoperability including European technologies such as a common European Rail Transport Management System and train control system (ERTMS/ ETCS), and establishing a common approach to European rail safety.

\* Jan Scherp is principal administrator at the European Commission, Brussels. The author takes sole responsibility for the views expressed.

In the following sections this paper will briefly describe the European Community's approach to (de-) regulation of the railway sector. It will also highlight recent rail market developments which have been triggered by the regulatory changes and review the key European Community's regulatory provisions for the rail sector in the light of these developments and policy objectives. The paper finishes with addressing the future challenges that the European railway sector must face in order to successfully revitalise the European railway sector.

### **The European Community's approach to regulatory reform in order to revitalise the railway sector**

A cornerstone of regulatory reform in the European Union is the opening and integration of the formerly closed monopoly railway markets. In freight transport, an open access approach was selected which enabled competition "on the tracks". In passenger transport the approach proposed by the European Commission was effectively two-pronged: open access for international services and regulated competition, for instance, for urban services through the tendering of franchises or public service contracts ("competition for the tracks"). Rail operations on the European infrastructure are a mix of passenger and freight traffic. In order to facilitate the creation of effective network opening and competition, the fragmented and multiple nationally integrated rail systems required infrastructure management to be separated from transport service provision. As a minimum, this separation must be done for the essential functions so as to ensure non-discriminatory network access such as capacity allocation and setting of track access charges. This fundamental choice should be kept in mind when compared to the North American or the Japanese approach. The North American rail system is predominantly freight oriented with very little passenger operations. A small number of integrated rail freight operators provide national services on often parallel, competing infrastructure and negotiate network access between each other, if required. In Japan, the majority of rail traffic is passenger transport provided by integrated regional monopoly operators.

In the European Union, the first milestones to be achieved to fully open and integrate the rail markets were the rail interoperability and rail infrastructure package directives of 2001. The three directives of the infrastructure package ("the first railway package"), 2001/12/EC (on the development of European railways amending Directive 91/440/EEC), 2001/13/

EC (on railway licensing amending Directive 95/18/EC) and 2001/14/EC (on capacity allocation, railway infrastructure charging and safety certification), defined the access rights to use rail infrastructure for international freight services, the various conditions railway firms must fulfil to be able to benefit from the access rights, the independence of functions essential for ensuring non-discriminatory access and the possibilities of appeal that the market actors should have. The directive on interoperability of conventional rail systems (Directive 2001/16/EC) describes, similarly to the high-speed rail directive (Directive 96/48/EC), a process of technical harmonisation of the railway based on Technical Specifications for Interoperability.

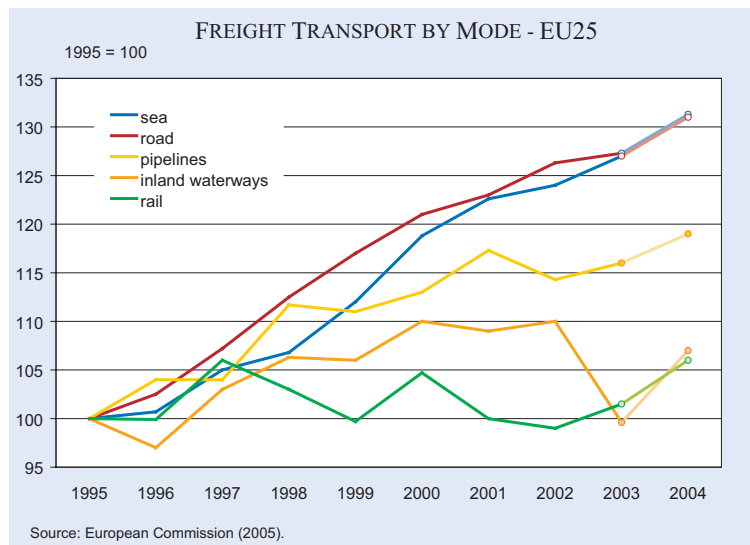
The regulatory reform continued through the adoption of a second legislative railway package in April 2004. The second package provided for full open access for all kinds of rail freight services, a common approach to European rail safety, extending the scope of the interoperability directives and the setting up of a European Railway Agency in Valenciennes (France). The ERA will drive forward the technical implementation of the EU safety and interoperability approach. In March 2004, the European Commission made proposals for the market opening of international rail passenger services, enhancing rail passenger rights and completing the interoperability legislation by extending it to a common licensing regime for train crews.

### **Recent rail market developments – can the policy targets be met?**

It is certainly too early to draw any firm conclusions on the precise effects of the regulatory changes triggered by the EU rail directives of 2001, as most member states did not transpose them into national legislation until 2003 or later. However, some member states had already started to open up their rail markets in the early 1990s and thus some preliminary trends at least for rail freight can be identified.

The traffic performance of EU rail passenger transport has hardly changed in absolute terms between 2000 and 2003 in the enlarged European Union of 25 member states. The traffic performance remained around 345 billion passenger km per year. In the old member states (EU15) the share of high-speed rail passenger services rose from 19.4 percent (2000) to 23.1 percent in 2003. Whereas in 2004 the passenger transport in the EU15 rose by 0.5 percent, in the new

Figure 1



member states it dropped by ca. 7 percent reflecting the increasing importance of car driving in these rapidly developing economies. The traffic performance is best in countries where competition is the most developed such as the UK and Sweden, as well as France due to their development of high-speed services.

In rail freight transport the sought after modal shift has not yet been achieved. Figure 1 shows that the market has remained essentially flat since the mid-1990s. Traffic performance picked up recently, however. Between 2003 and 2004 it rose by 5.8 percent in the old member states (EU15) and by 4.4 percent in the enlarged Union (EU25).

Employment in railway firms in the EU15 dropped from ca. 1.3 million in 1990 to ca. 770,000 in 2000 due to declining market shares and restructuring of the railway sector. This included the subcontracting and outsourcing of certain rail related activities such as maintenance. In the enlarged EU25 the current employment level amounts to around 1.1 million of which roughly 65,000 are employed in new entrant and private railway undertakings (estimation based on European Foundation 2005). The pace of employment reduction has slowed down considerably in the last two to three years. Job losses in the incumbent railway firms have partly been compensated for by the creation of new jobs in new railway businesses.

Rail freight transport is increasingly becoming a European business. Currently, some 50 percent of rail freight services in the EU are international (imports, exports or transit). The share of international services varies greatly between smaller countries, where it is

higher, and bigger member states, where it is lower due to the relatively higher importance of the domestic freight market. On some major European rail corridors such as the one between Rotterdam and Genoa, traffic performance has increased in recent years from around 5 percent to 10 percent (CER 2005). This growth has been realised mainly due to block train/shuttle train activities where the new entry of railway firms has so far been the strongest.

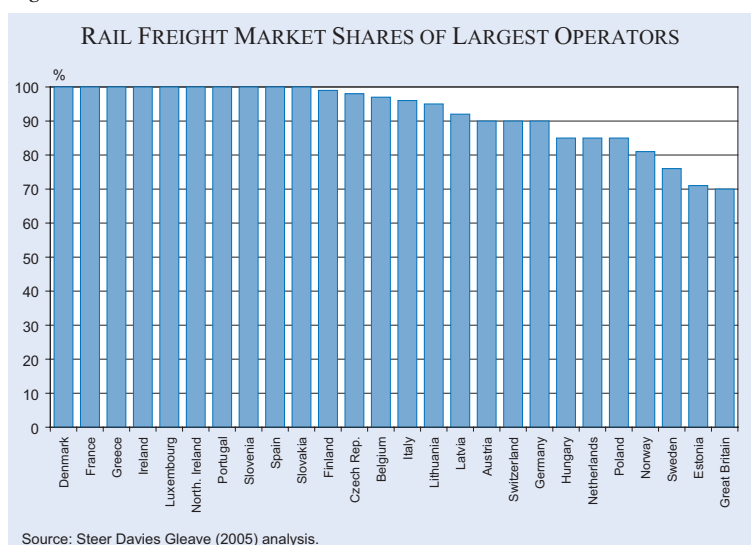
Rail freight price is strongly influenced by the prices used in related competing markets, such as road haulage. The evidence on price trends based on national data is not conclusive, although some anecdotal evidence points to a price-reducing effect on more open markets. In 2004, prices fell in Germany by ca. 9 percent and on the Rotterdam-Genoa corridor by 15 to 20 percent over the last three years. On more closed markets prices appear to have increased, for example, by an average of just over 2 percent annually in Belgium and in France by 8.5 and 18 percent in 2003 and 2004, respectively, compared with typical past increases of only 2-4 percent (Steer Davies Gleave 2005).

The opening of rail freight markets, based either on Community legislation or national initiatives, has resulted in increased market entry in recent years, although at a very modest scale. In 2003, new entrants had an estimated market share of around 3 to 4 percent in terms of turnover in the Community of 25.<sup>1</sup> Market entry was particularly strong in Germany, Sweden, the UK, the Netherlands, Italy, Poland and the Czech Republic. Many new entrants are small and operate in only one member state. An analysis of national markets shows that there remains a clearly dominant operator on every network (see Figure 2).

Since market opening a certain trend towards market concentration through mergers and acquisitions can be observed. Some rail freight operators have developed a European business strategy and positioned themselves in several national markets, for example the Railion group (in Germany, Netherlands, Den-

<sup>1</sup> Personal communication of the Secretary General of the European Rail Freight Association (ERFA) as representative of the interests of the new entry rail freight railway undertakings.

Figure 2



mark and Italy), Trenitalia in Italy and Germany (through acquiring a majority share in the private German company TX Logistik) and the Swiss SBB Cargo by setting up subsidiaries in Germany and Italy. Should the recent trend of historical operators entering into competition with each other be confirmed then this would have a significant impact on the level of competition on the European rail freight market. For instance, since December 2003, Railion and its Swiss partner BLS Cargo compete with SBB Cargo and its subsidiaries on the North-South corridor through Switzerland.

By entering into new alliances, new entrants are now in a position to provide competitive international services and to compete with national incumbents for such services. The “European Bulls” alliance set up by five new entrants in January 2005 is one example for this new trend. Faced with the high market share and relatively strong capital base of the national incumbents, new entrants often fight an uphill battle to establish a significant position on the market. They clearly require favourable market regulation and transparent, non-discriminatory framework conditions to succeed.

### Reviewing the regulatory framework

The directives of the first railway package, which form the basic regulatory framework for rail transport in the Community, have been in place in most EU member states for more than two years. However, how the national regulatory framework functions and the administrative efficiency in implementing it vary from country to country. A key question

of European regulation is whether the framework put into place is helping to achieve the desired political objectives. The following section will review the four key provisions of the Community legal framework in more detail: transparency as well as the independence of essential infrastructure management functions for non-discriminatory access (e.g. capacity allocation, setting of track access charges) for transport provision, access to rail service facilities, rail infrastructure charging, and the functioning of the regulatory body.

### *Separation of infrastructure management from rail service provision vs. integration*

The directives of the rail infrastructure package contain a number of provisions requiring increased transparency on how rail activities are carried out and on their funding. This is in order to ensure a non-discriminatory market access for all railway firms and thus an effective market opening. Basically, three levels of separation are required:

- Accounting separation between infrastructure management and rail transport provision,
- Separation of accounts between rail passengers and rail freight activities disallowing any transfer of public compensation payments for public passenger services to freight transport activities,
- Independence of essential functions (e.g. infrastructure charging, capacity allocation, licensing and safety certification) from rail service operations.

Accounting separation between infrastructure management and rail transport provision, which was due already under Directive 91/440/EEC, has been implemented in almost all member states. Accounting separation between freight and passenger transport activities is gradually being introduced, although some member states need to make further progress. In order to ensure the independence of essential functions such as track access charging and train path allocation member states have put different institutional structures in place. Basically, there are three major variants: institutionally fully independent infrastructure managers, such as in the UK, Sweden and the Netherlands; independent infrastructure managers or allocation bodies that rely to a certain extent on the

expertise and staff of the historical railway firms such as in France, and legally and organisationally independent infrastructure managers that are part of a railway holding structure, as in Germany, Poland, Italy and Belgium. There has been an intensive academic and political debate on the merits and disadvantages of separation. The outcome is still inconclusive in the absence of reliable, empirical evidence in the European context. However, it is clear that some minimum requirements must be fulfilled by infrastructure managers in order to ensure non-discriminatory network access.

Bodies or firms that are in charge of train path allocation and rail infrastructure charging must ensure that their institutional and procedural arrangements are inherently capable of guaranteeing the required independence. Criteria such as not having common board members with any railway firm, not having offices in the same building, being statutorily or contractually independent from holding companies or railway firms have to be fulfilled. Additionally, the arrangements must prove in practice that they assure the required independence, e.g. through the absence of substantiated complaints. If the results of this ongoing assessment are not satisfactory, legislative initiatives need to be considered for ensuring full independence, for instance, through institutional separation.

*Non-discriminatory access to and charges for rail related services*

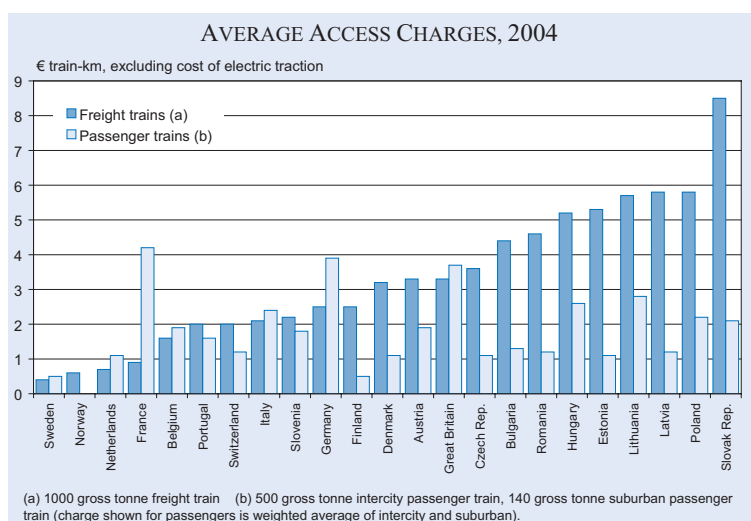
Railway firms will not be able to make use of railway infrastructure unless they also have access to appropriately priced rail-related service facilities, such as terminals, maintenance workshops, shunting yards, refuelling points and driver training. There have been complaints by new entrants about the difficulties of obtaining access to such facilities. New entrants sometimes face incumbents who are effectively able to deny access to the facilities. This is because the national incumbent either retains strong ties to the infrastructure manager due to their historical relationship, or controls access to terminals by means of management or ownership. As to the charging framework for such rail-related services it is often un-

clear what is being charged for and whether the charges are fair and accurately reflect the costs incurred. In some countries the charges are not defined and are thus not transparent in particular for new entrants. Envisaged for 2006 is a Community-wide, thorough investigation of the access situation and charges for the use of rail-related services as well as the future capacities of service facilities with a view to proposing measures to lower these market entry barriers and to ensure dynamic development of rail services.

*Charging for the efficient use of infrastructure*

The right structure and level of rail infrastructure charges is a key to incentivising railway firms and infrastructure managers to efficiently use the rail network. The track access charge also contributes to the financing of the maintenance and development of the infrastructure. The EU Directive 2001/14/EC defines as a basic principle that the track access charge should be set at the cost that is directly incurred as a result of operating the train service. It also allows for adding other cost elements such as mark-ups which reflect the fixed cost elements of operating the network or reservation charges ensuring that operators use the paths they have requested. The general charging principles have been implemented very differently by member states leading to greatly varying structures and levels of charges as well as cost recovery rates across the Community (see Figures 3 and 4). This situation can lead to confusing or even conflicting incentives for international rail transport.

**Figure 3**

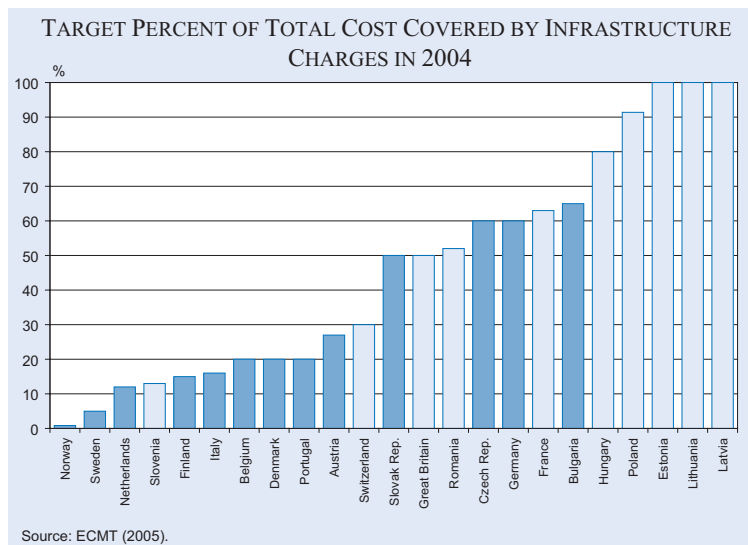


Baltic freight trains are much larger than elsewhere. Baltic access charges are not directly comparable with those in other countries and have been adjusted here. In Estonia, for example, a typical 3 145 tonne train is charged € 11 per train-km.

Source: ECMT (2005).



Figure 4



High levels of cost recovery gained as a result of the access charge can reduce the competitiveness in particular of rail freight traffic relative to other modes of transport. To avoid undesirable modal shift effects infrastructure managers should ensure that track access charges on major European corridors are set at levels that are consistent with the charges to be paid for alternative infrastructure, such as roads. Furthermore, if the infrastructure manager is entitled to recover all costs through access charges, there may be a reduced incentive to improve operational efficiency. It may be better to have a lower cost recovery objective focused only on the recovery of variable costs, which incentivises the infrastructure manager to control operating costs. This could be combined with a well-defined service contract to maintain and renew the network to a certain level of quality with key targets for efficiency and performance improvements. Designing such infrastructure service contracts is a delicate task. Whilst incentivising infrastructure managers to operate efficiently they should also ensure their financial stability by ensuring predictable direct financial support from the State or other sources.

Performance regimes effectively incentivising infrastructure managers and railway firms to avoid disruptions of the network (e.g. through applying a financial bonus-malus scheme) have been introduced only in a few member states such as the UK and Denmark. In view of ensuring efficient use of the rail infrastructure, including that for international services, member states should introduce incentivising performance regimes in line with implementation guidelines at the European level. This could be further developed by RailNetEurope, the Vienna based technical co-operation platform of rail infrastructure managers.

### *Effective regulation for enhanced competition*

The regulatory body, to be set up according to Art 30 of Directive 2001/14/EC, has a key role in ensuring non-discriminatory access to rail infrastructure and service facilities as well as overseeing the development of competition on the rail service market. In some member states, the rail regulator is fully operational whereas in a few others it has not yet been set up or it is not yet completely operational. The existence of credible regulatory bodies which have the appropriate administrative ca-

capacity to act as an appeal body is necessary in order to actively promote market entry. Stakeholders often claim that the regulatory body should be fully independent, not only of infrastructure managers and railway firms but also of the State. Currently, in many cases the regulatory body either reports to a public authority or its functions are executed by a ministerial service. As the national rail operator and the infrastructure manager are generally owned by the State there may be a potential conflict of interest. Hence the merits of complete independence of the regulatory body from the State should be seriously evaluated.

### **Concluding remarks and outlook into the future**

Although some railway firms have been successfully restructured through the establishment of commercially oriented rail services providers, the objective of market integration has not yet been fully achieved. In some member states significant market entry has occurred leading to a certain degree of competition on domestic rail transport markets and on a few European rail corridors such as the north-south corridors across the Alps. The targeted increases of rail traffic performance have not yet been reached, although a certain upward trend has been realised in rail freight from 2003 to 2004 and, in particular, in countries with open rail freight markets. In order to more effectively integrate the rail market and facilitate competition in the future, further efforts including regulatory initiatives in this area are required. More effective promotion of market entry requires independent management of infrastructure and service facilities, coherent and efficient price signals for



using this infrastructure as well as powerful and independent regulatory bodies.

Furthermore, in order to ensure a prosperous future for the railway sector, railway business must become simpler and less costly. For instance, lengthy and costly procedures to obtain a railway licence and a safety certificate are entry barriers that must be lifted in the future. Mutual recognition of testing routines for rolling stock is still problematic and non-transparent and diverging insurance requirements as well as the absence of a truly common European rail insurance market render the preparation of international freight services a difficult task. The gradual implementation of the safety directive and the growing role of the European Railway Agency (ERA) set up in 2005 are expected to help overcome some of these problems in the coming years. The ERA will act as a future network integrator through facilitating mutual exchange between the competent national authorities and enhancing an EU wide build-up of expertise in the field of interoperability and safety.

Likewise, more co-ordination is needed for the future development of an integrated European rail infrastructure that meets the needs of the market. The loosening of the formerly close relationship between rail services provision and infrastructure management requires the putting into place of appropriate mechanisms to co-ordinate the investment plans for infrastructure development with the future development of rail transport services, as expected by the various railway firms operating on the network. EU-wide co-ordination also needs suitable priority allocation rules applied coherently along European corridors in case of network saturation.

Currently, railway firms planning to provide seamless cross-border services are still obliged to use expensive multi-system locomotives to cater, for instance, for the various national train protection and traction energy systems. The technical market integration based on the implementation of the Technical Specifications for Interoperability (TSI) is likely to take many more years. The long investment cycles in the railway sector tend to delay the convergence of the national technical systems in view of establishing full interoperability. It is therefore important to start to invest in interoperable technologies when the old national systems have reached the end of their lifetime or where payback periods are sufficiently short to justify the investment. Typical examples are train traffic management and control sys-

tems in particular in new member states of the Community, and electronic data exchange systems, for instance, to enable cross-border tracking and tracing, and fleet management applications. The co-ordinated deployment of the European Rail Traffic Management System ERTMS/ETCS and the setting up of a European platform for electronic data exchange based on common, European specifications are challenges of the immediate future. The coherent implementation of such systems requires close co-operation between railway firms, infrastructure managers and public authorities.

The future economic and financial viability of the railway sector in particular in the new member states of the European Union is a major challenge. Although the financial restructuring of many historical railway firms in the old member states in the mid-1990s led to a reduction of their debt from € 130 billion to € 100 billion (NERA 2003), the growing shortage of public funds for the development and maintenance of the rail infrastructure constitutes a potential threat for the future development of the railway systems. These problems are particularly severe in many new member states. The finances of the railway systems are often completely out of balance, resulting in a vicious circle of insufficient investment and maintenance of the infrastructure leading to substandard transport services, falling demand for services, rising request for public support for socially desirable transport services and lower amounts of public funds available for infrastructure expenditure, etc. Hard choices need to be made in terms of defining a sustainable network size, and the solutions for its funding need to be found in order to put the railway systems on a sound financial footing.

The European Community intends to continue its support of the development of rail infrastructure and modernisation of rolling stock through its various funding mechanisms, such as the Trans-European Network (TEN), structural and cohesion funds. For the financial period 2007–13 the European Commission proposed that a four-fold increase of the budgetary means up to € 20 billion be made available for the TEN-Transport development.

The assessment of the European railway sector shows that there can be a prosperous future if the major challenges can be taken up successfully. These challenges are ensuring open and competitive rail service markets, bringing down market entry barriers and systems costs, making swift progress towards

a European network integration and successful delivery of a sustainable financial restructuring.

## References

CER (2005), *Recent Trends in the International Rail Freight Market: A Case Study of Rotterdam-Genoa*, October.

ECMT (European Conference of Ministers of Transport, 2005), *Railway Reform and Charges for the Use of Infrastructure-Conclusions and Recommendations*, CEMT/CS(2005)9 of 4 April 2005, Paris.

European Commission (1996), *White Paper on "a Strategy for Revitalising the Community's Railways"*, COM(96) 421 final, Brussels, 30 July 1996.

European Commission (2001), *White Paper on "European Transport Policy for 2010: Time to Decide"*, COM(2001) 370 final, Brussels, 12 September 2001.

European Commission (2005), *Consultation Document of the European Commission Directorate General for Energy and Transport on the Mid-term Review of the Commission's 2001 Transport White Paper*, [http://europa.eu.int/comm/energy\\_transport/white\\_paper\\_transport\\_revision/lb\\_consultation\\_questionnaire\\_en](http://europa.eu.int/comm/energy_transport/white_paper_transport_revision/lb_consultation_questionnaire_en).

European Foundation for the Improvement of Living and Working Conditions (2005), *Employment and Working Conditions in the European Railway Transport Sector*, draft report, Dublin.

NERA(2003), *Study of the Financing and Public Budget Contributions to Railways*, final report for the European Commission, London.

Steer Davies Gleave (2005), *RAILIMPLEMENT – Implementation of EU Directives 2001/12/EC, 2001/13/EC and 2001/14/EC*, draft final report compiled for the European Commission – Directorate General for Energy and Transport, November 2005, Bologna.

## INEQUALITY ACROSS THE GENERATIONS IN NORTH AMERICA AND EUROPE

MILES CORAK\*

In thinking about the welfare state, about past accomplishments and future challenges, I would like to begin by borrowing a few pages from the work of Amartya Sen, the 1998 Nobel laureate in economics. Sen's thinking is neatly summarized in his popular book *Development as Freedom*, which in the first instance is directed to issues in development economics, but has broader implications and lessons for social policy in rich as well as less rich countries. Sen argues that the major goal and major means of development should be freedom. Society should offer its citizens the freedom to lead the lives they choose to value.

This idea has in fact been an important undercurrent in the development of the welfare state in North America and Europe, from its earliest days in the 1800s through to the end of the Second World War. Then the emphasis was on the importance of full employment and the need for economic security for those excluded from jobs because of business cycle recessions, because of sickness, or because of the physical limitations of old age. But there has also been an emphasis on active measures involving investments in human capital, first through more and more years of compulsory schooling, then to the broadening of access to higher education and more recently in some countries to early childhood development.

Sen's thinking asks policy makers to focus on the capabilities of citizens and to remove barriers to full participation in society, thereby allowing them to make use of their talents and develop their full potential. He frequently cites the quest for freedom from poverty, adequate health care and education as examples of specific social policies affording citizens these capabilities.

I choose this as my starting point to draw what I think is an obvious link, though one not directly made by Sen or others, to the circumstances of children. There

is in the case of this particular group, I think, a connection between on the one hand freedoms and capabilities, and on the other hand direct measures of these concepts as well as policy levers. Freedom means that children can become all that they can be. In some large sense this means that they are free to realize their full talents, and that their outcomes in life are not tied to their family backgrounds. In other words, for children this means equality of opportunity.

As an economist I am inclined to think about these issues in terms of labour market outcomes, though obviously the issues are broader than just earnings and incomes. The subject of inequality is central to any discussion of the role of the welfare state, but in this essay I will ask you to think not just about inequalities in the here and now but also about how these inequalities play out across generations.

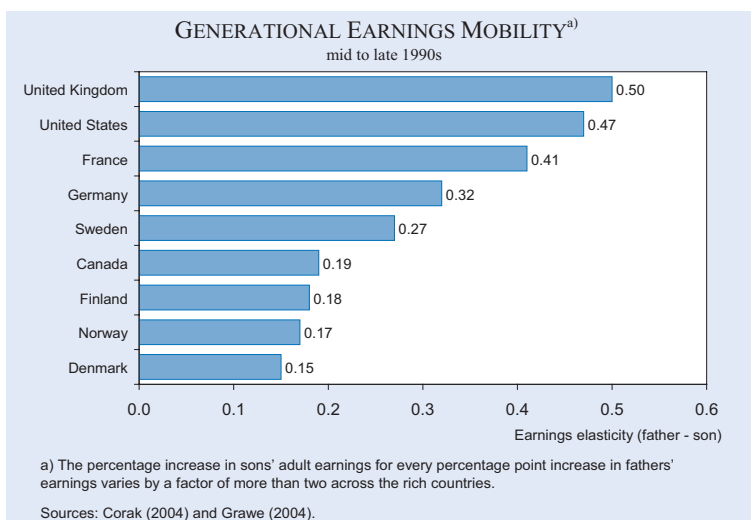
There is a sense in which a high level of income inequality need not necessarily reflect, in Sen's words, an "unfreedom" and would in fact be tolerated by even the poorest in society. This would be the case if there were flux and mobility across the generations, if, in spite of current circumstances, the children of the poor were as likely to grow up to be high-income adults as the children of the rich. Similarly, the same high level of inequality in the here and now may have very different implications for social cohesion and individual welfare when there is very little generational income mobility, when child outcomes are strongly tied to the circumstances of the families in which they were raised.

One direct measure of generational income mobility is the strength of the link between an individual's earnings and his or her parents' earnings when they were raising their family. This tie will determine the income advantage, relative to the average family, that higher income parents pass on to their children. Or for that matter the income disadvantage that lower income parents pass on.

The percentage increase in a child's adult income for every percentage point increase in income of his or her parents is presented in Figure 1 for a number of OECD countries. This information is based upon the comparative research by Nathan Grawe (2004) and a broader synthesis of the literature by myself in Corak (2004). The higher this statistic, the lower the degree of generational mobility. These findings suggest that first there is a good deal of variation across the rich countries – by at least a factor of two – in the degree

\* Miles Corak is director of Family and Labour Studies at Statistics Canada. He is the editor of *Generational Income Mobility in North America and Europe* published by Cambridge University Press, on which this article is based.

Figure 1



to which an earnings advantage is passed to children. Second, in no country is the inherited parental advantage much lower than one-fifth. The United States, the United Kingdom and to a slightly lesser extent France stand out as being the least generationally mobile societies, with every 10 percent increase in parental income implying that the grown-up child incomes will be 4–5 percent higher. At the other extreme are Denmark, Norway, Finland, and Canada where the relationship is more than half as weak.

In the United States households with children under the age of 18 at the top income quintile had 12 times as much money as those at the bottom quintile. The information in Figure 1 can be used to translate this ratio into the economic advantage a child from the higher-income family can expect to have in the next generation over one from the lower income family. The 0.47 statistic reported for this country implies that the adult income of someone born to a family at the top would be almost three and one third times higher than someone born to a family at the bottom. With a value in the order of 0.2, as for example in Canada, this income advantage would still have been significant, but at less than 66 percent much smaller.

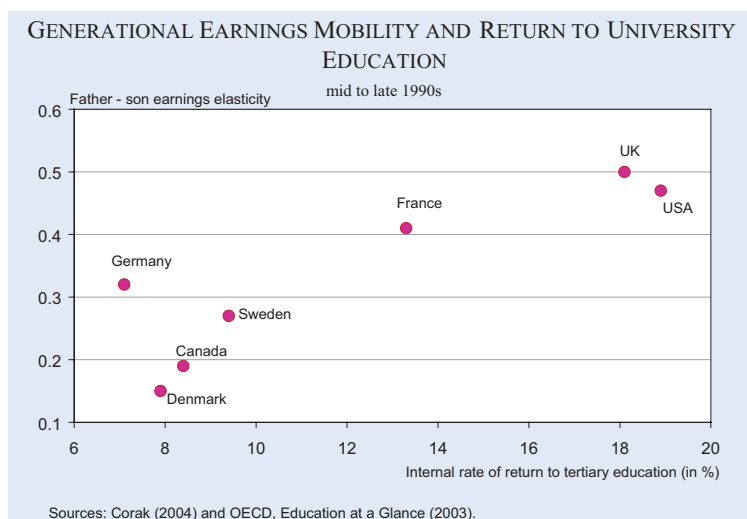
But what exactly does this statistic mean? The information in Figure 1 is purely descriptive and on its own tells us nothing about the reasons for the degree of generational mobility or for the differences among countries. Effective policy intervention requires not

just an awareness of raw correlations, but also an understanding of causal processes. I would like to focus on two of several mechanisms that underpin these generational mobility figures and that open up distinct opportunities for the conduct of public policy. The first has to do with how labour markets work, and the second with the relative benefits of public policy.

By expanding upon the standard theoretical models used by economists to study generational earnings mobility Gary Solon (2004)

argues that more labour market inequality implies less generational mobility. His analysis suggests that in an economy emphasizing human capital as the basis for sustained growth – one in which the economic returns to education have a tendency to rise – more challenges will be placed in the way of generational mobility. Indeed, one important determinant of the degree of earnings inequality is the return to higher education. Figure 2 shows that a higher return to university education is associated with tighter links between father and son outcomes. The Figure presents a scatter plot between the estimates of generational income mobility from Figure 1, and the private rate of return to university education (relative to secondary education). For these countries there is a clear positive relationship. The three countries with rates of return higher than 10 percent – the United States at 18.9 percent, the United Kingdom at 18.1 percent, and France at 13.3 percent – are the countries with the least generational earnings mobility.

Figure 2



Jo Blanden, Alissa Goodman, Paul Gregg and Stephen Machin (2004) also use the framework offered by Solon to examine variations in the degree of generational mobility over time within a specific country, the United Kingdom. They find that the degree of mobility actually declined between a cohort of young people growing up in the 1960s and 1970s and another growing up a decade later. The authors explain this by noting that this period was one of widening wage and income disparities. But they also note that while the educational attainments of the young increased significantly over this period, they did so in a way that was of relatively greater benefit to those with higher-income parents.

In addition to shedding light on developments over time in one country this suggests that there could well be a number of possible explanations for the cross-country patterns observed in Figure 2. In particular it may be that countries with higher rates of return also have very different structures and policies in place in terms of access to higher education, since one of the reasons for an elevated rate of return is the presence of restrictions in the supply of university graduates. In this sense it may also be that a higher rate of return reflects inequality of opportunity. And, further, higher rates of return do change the incentive for parents, leading those who value education most to invest even more in their children.

This explanation focuses on the structure of rewards and private investments in children, but it also raises the question of differences in opportunity, that is, the extent to which children from higher income families are more likely to capitalize on the rewards available to higher education. Thus, the second factor likely to explain cross-country differences in generational mobility has to do with the opportunities children are afforded and the nature of public investment in them.

Susan Mayer and Leonard Lopoo (2004) point out that increases in “progressive” investments – those of relatively more benefit to the less well off – will loosen the link between parent and child outcomes. Traditionally, this was seen as an important aspect of public schooling. Societies differ a good deal in their levels of spending on education. This is illustrated in

Figure 3

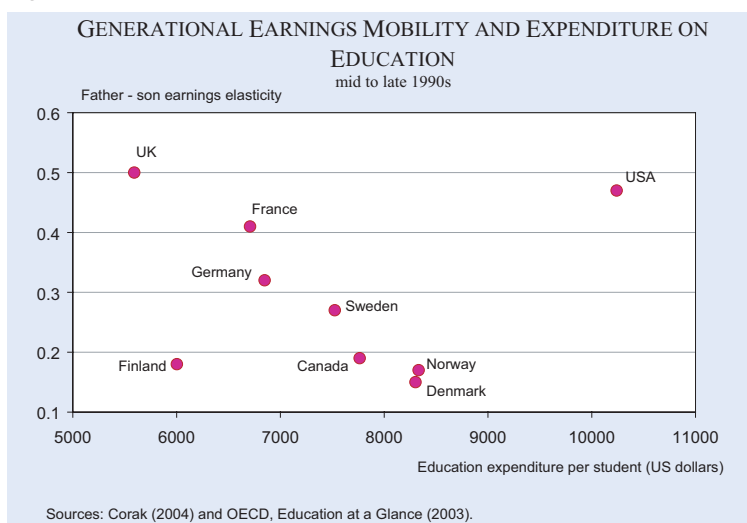


Figure 3, which plots the information from Figure 1 against the education expenditure per student.

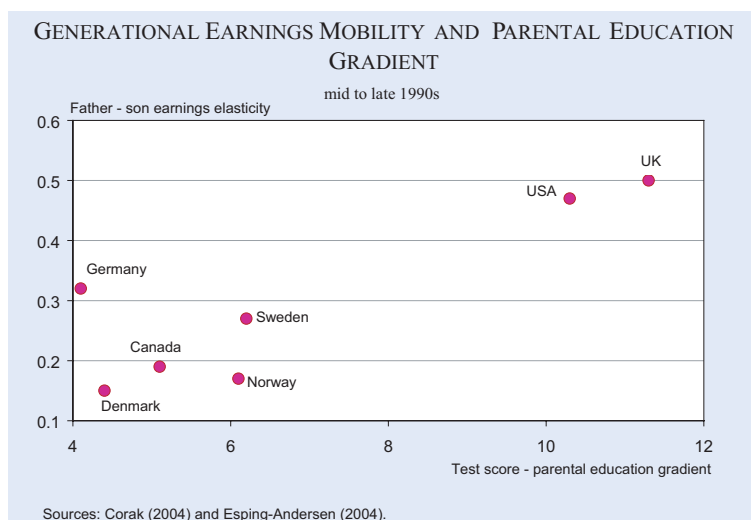
More spending per student is generally associated with more generational mobility. But the relationship between spending on education and the degree of generational mobility is not as simple as that. For instance, the United States is a clear outlier, with the highest levels of spending but also one of the strongest tie between fathers’ and children’s earnings. On the other hand, the United Kingdom and Finland both spend relatively low and similar amounts per student but are characterized by very different degrees of generational mobility.

The important point from these results is that not only the overall level of public spending on education matters, but also *how* the money is spent. In other words, the way education systems are structured and how the cognitive capacities of children are developed to allow them to take advantage of whatever opportunities are made available is important.

The best way to understand these patterns is to recognize the extent to which public investments are of relatively more benefit to the disadvantaged. One possible measure of this is how tightly related the abilities and skills of children are to the educational levels of their parents. Esping-Andersen (2004) argues, for example, that if literacy and numeracy skills in adulthood are strongly correlated with parental educational levels, this suggests that spending on education has done little to level out relative advantages and disadvantages that are based in the home and possibly that public investments have not been terribly “progressive”.



Figure 4



This information is available for seven of the countries highlighted in Figure 1, and the relationship with generational earnings mobility is illustrated in Figure 4. In these countries there is a very strong positive relationship between the numeracy/literacy skills of adult children and the educational levels of their parents, with Germany being the only significant outlier.

This pattern reflects the inequality of private and public investment in children. If the degree of income inequality is higher in one country than another it might be reasonable to expect that not only will the rewards of a given level of investment be higher, but also that the level and distribution of investment in children will be different. A country with more income inequality might also have more inequality in the investment that rich and poor parents can make in their children, and hence a lower degree of generational income mobility.

But the extent to which this is so will also depend upon the degree to which children from less advantaged backgrounds disproportionately benefit from public programs. Under certain circumstances universal government programs can reduce the investment gap between rich and poor children. If the first dollar of investment creates the greatest increase in the well-being of the child, then when government invests equally in all children, poor children are likely to gain more than affluent children. The information in Figure 4, for example, suggests that the UK and the US get relatively little punch in terms of higher generational mobility for every dollar spent on education because the structure of their educational systems reinforces the relative advantages and

disadvantages children get from their home environment rather than levelling them out.

Indeed, the great promise of government investment in the expansion of universal education and increased access to higher education during the postwar period was that it would give children from relatively disadvantage backgrounds an extra push and put them on a par with their fellow students. Figure 4 suggests that there are still very large differences in the extent to which this has been done in the welfare states of

the rich countries, and as such offers an important hint as to why countries differ in the degree to which economic advantage is passed on between parent and child.

In sum, the generational mobility of earnings in the Nordic countries and Canada is higher than other rich countries because first, labour market inequalities and the returns to education are relatively lower, and second, the mix of public and private investments in children has been relatively progressive. But this story reflects the situation of the last 30 years or so, a generation that was born in the 1960s, came of age and went to middle school and university in the 1970s and 1980s, and found its place in the labour market of the 1990s. It cannot be uncritically transposed to today's newborns and elementary school children, who will be attending the colleges and universities of the 2020s and working in the labour market of the 2030s.

One important future challenge concerns access to education. In some countries – the United Kingdom and Canada for example – a climate of higher tuition fees is on the horizon, and it is also likely that the education systems in these countries may also witness more decentralization in the way fees are set. Fees on average will likely continue to rise, but they may also vary a good deal more – between institutions and also fields of study. A much more differentiated post-secondary system is in the offing.

In this context there is a growing concern about access to university education. And while there may be a need for more public funding and for reconsideration of the structure of supports to students, particularly to those from low-income backgrounds, the issues of ac-

cess are more than just financial. Access to higher education is often discussed in very broad terms, that is, whether students are more or less likely to attend university according to their family income. In the coming years the significance of access to higher education may also be more narrowly framed to refer to access to particular institutions or fields of study. Generational mobility will also be influenced by the nature of universities' student selection criteria.

If children from higher income families are more likely to have the skills to gain admittance to those areas most highly rewarded in the labour market, a rise in admission standards and the development of other selection criteria may lead to stronger links between family background and post-secondary participation in particular institutions or fields of study. In this sense it is important for policy-makers to understand the non-financial barriers to accessing higher education, particularly circumstances earlier in the lives of young people that help them continue their education after middle school.

This is why Esping-Andersen (2004) argues that concerns about generational mobility may lead policy-makers to focus increasingly on the family, particularly its role in the development of children's cognitive abilities. Cognitive and social skills are an important determinant of long-run earnings; they are only loosely correlated with educational attainment; and cognitive performance is more closely related to the "cultural capital", or to put it more broadly, the "parenting style" of the family, than it is to its material wealth. In fact, the kinds of parental investments that are decisive are not the monetary kind. The inheritance of education, occupation and income is influenced in the first instance by the impact parents have on a child's cognitive performance, and – as Figure 4 illustrates – societies leveling the playing field with respect to these circumstances have had the most success in promoting generational mobility.

This is also why even in the Nordic countries and Canada as many as one-third of children from low income backgrounds grow up to be low income adults, in spite of the loose overall tie between the incomes of children and parents. On average these societies are very mobile across the generations, but a significant proportion of children fall through the cracks in this fabric.

Early childhood initiatives may play an important role in determining the degree of generational mo-

bility. If this is the case social policy may increasingly be called upon to equalize the impact families have on children's skills, beliefs and motivation. The impact of early childhood initiatives on generational mobility will depend upon the extent to which they prove to be effective and of relatively more benefit to children in disadvantaged families. But John E. Roemer (2004) offers a cautionary note for policy makers concerned with equality of opportunity. Does equality of opportunity imply that there should be no correlation in incomes across the generations? Should in other words governments set the absence of any correlation in generational incomes as a target to guide policy?

Roemer answers these questions by first noting that equality of opportunity implies that inequities of outcome are indefensible when they are due to differential circumstances, but also by noting that societies and parents influence their children through a hierarchy of circumstances. If we are to understand what equality of opportunity means and how it can be influenced, we have to know what these circumstances are. In Roemer's view parents influence their children through a hierarchy of circumstances: through social connections that facilitate access to education and jobs; through family culture and investments that influence beliefs and skills; through the genetic transmission of ability and through the formation of preferences and motivations. These are the successively broader fields – each corresponding to a successively broader definition of equality of opportunity – which policy makers could potentially seek to level.

Roemer makes explicit that equating equality of opportunity with complete generational mobility – with no statistical tie between parent and child earnings – implies that not only should the influence of social connections and also of family culture and investment be eliminated, but so should the genetic transmission of ability and the influence of family on the formation of preferences and goals among children. He argues that this is "a view that only a fraction of those who consider the issue would, upon reflection, endorse".

This is a cautionary note: to eliminate entirely the income advantage that is passed from parents to children would require a degree of intervention into the lives of children and families that the majority in most societies would find untenable. The degree to which the parental income advantage passed on to

children is consistent with equality of opportunity is not self-apparent. It requires a definition of the circumstances unacceptable as sources of labour market success, an understanding of the effectiveness of policy interventions, and recognition of the trade-offs between the gains in eliminating them and the losses in terms of other measures of welfare.

The capacity of children to become self-sufficient and successful adults is compromised not only by monetary poverty, but by poverty of experience, influence and expectation. This argument calls for broader thinking on the mechanisms and causes of generational mobility, and the extent to which it draws governments into broader areas of social and family policy will depend upon societal values.

But what is clear is that the postwar agenda of offering increased access to higher and higher levels of schooling seems to have reached its limit. If the rich societies wish to continue to promote equality of opportunity, then they will need to invest more in children earlier in their lives to ensure that they have the skills and the opportunities to succeed in the labour market. This shift in direction may involve a conception of equality of opportunity that has different degrees of support, because it requires that public policy influences the impact that families have on their children's skills, beliefs and motivation.

In this sense, whether the welfare states of the rich countries are able to offer effective programs of relatively more benefit to the least advantaged is the major challenge determining whether the next generation will enjoy the same degree of opportunity, or – to use Sen's word – freedom, as the last generation.

## References

- Blanden, J., A. Goodman, P. Gregg and S. Machin (2004), "Changes in Intergenerational Mobility in Britain", in M. Corak, ed., *Generational Income Mobility in North America and Europe*, Cambridge University Press, Cambridge.
- Corak, M. (2004), "Do Poor Children Become Poor Adults? Lessons for Public Policy from a Cross Country Comparison of Generational Earnings Mobility", unpublished paper presented to the "Colloque sur le devenir des enfants de familles défavorisées en France", Paris, and to the workshop on "The Welfare State in an Intertemporal Perspective", University of Siena.
- Esping-Andersen, G. (2004), "Unequal Opportunities and the Mechanisms of Social Inheritance", in M. Corak, ed., *Generational Income Mobility in North America and Europe*, Cambridge University Press, Cambridge.
- Grawe, N. D. (2004), "Intergenerational Mobility for Whom? The Experience of High-and Low-earning Sons in International Perspective", in M. Corak, ed., *Generational Income Mobility in North America and Europe*, Cambridge University Press, Cambridge.
- Mayer, S. E. and L. M. Lopoo (2004), "What Do Trends in the Intergenerational Economic Mobility of Sons and Daughters in the United States Mean?" in M. Corak, ed., *Generational Income Mobility in North America and Europe*, Cambridge University Press, Cambridge.
- OECD, Education at a Glance 2003, Paris 2003.
- Roemer, J. E. (2004), "Equal Opportunity and Intergenerational Mobility: Going beyond Intergenerational Income Transition Matrices", in M. Corak, ed., *Generational Income Mobility in North America and Europe*, Cambridge University Press, Cambridge.
- Sen, A. (1999), *Development and Freedom*, Random House, New York.
- Solon, G. (2004), "A Model of Intergenerational Mobility Variation over Time and Place", in M. Corak, ed., *Generational Income Mobility in North America and Europe*, Cambridge University Press, Cambridge.

## CONCEPTS AND MEASUREMENT OF LABOUR MARKET INSTITUTIONS

WOLFGANG OCHEL\*

### Introduction

Institutions have a decisive importance in determining how labour markets function. They place constraints on the individual behaviour of market participants and in so doing they may direct it in a certain direction. Institutions can reduce uncertainty and contribute to stabilising expectations. They may take the form of laws, ordinances, and legal precedents together with moral precepts, norms and customs.

The scientific analysis concerns itself, on the one hand, with the origin and evolution of labour market institutions. Ideally, they are created in order to reduce inefficiency due to market failure, in order to correct the distribution of earnings and in order to increase the security of income. Interest groups can also give rise to institutions (Blau and Kahn 1999). On the other hand, analysis of institutions is also concerned with their effects. By means of aggregate analysis of effects and microeconomic evaluation one seeks to estimate the influence exerted by labour market institutions on certain indicators or the influence they exert on market participants directly affected.

Whilst these two areas of analysis have received a great deal of attention, the labour market institutions themselves have been mostly neglected. But institutional arrangements must be captured adequately if one is to have an explanation for their origins and if their effects are to be understood. Capturing institutions requires that the domain regulated by an institution should be clearly defined and delimited. What is more, it is necessary to formulate a theoretical concept that can serve as the basis for capturing the institutions. Furthermore, the institutions must be investigated empirically. Apart from the assessment of institutional arrangements, qualitative information must be transformed into quantitative information.

And finally, it may prove to be necessary to aggregate individual indicators to a composite indicator.

The following article deals with capturing formal labour market institutions. Section 2 defines labour market institutions and presents the domains ruled by them. Section 3 provides a survey of the most important sets of data available for international comparisons. Section 4 presents the measurement concepts that are most prevalent and analyses the problems that arise in connection with the measurement of institutions. In sections 5 to 8, the institutional arrangements regarding employment protection, wage bargaining, the social security system as well as active labour market policy, and taxing labour are reviewed and the concepts on which they are based are analysed. Section 9 deals with the characteristics of institutions that up till now have been neglected when capturing labour market institutions. The summary in section 10 concludes the article.

### Definitions and domains

Labour market institutions are defined here by generally known rules that are designed to give structure to the recurring interactions in the labour market. If the enforcement of these rules involves recourse to the state's monopoly of the use of coercive force, then the institutions are referred to as formal institutions. Examples of such institutions are laws, ordinances and legal decisions. Institutions whose enforcement does not involve recourse to the state are referred to as informal institutions. Moral concepts, norms and customs are examples of such institutions (Voigt 2002).

Interdependencies exist between informal and formal institutions. Informal institutions may supplement formal institutions or may be a condition for their existence (Agell 1999). For instance, the generous safety net provided for by the Danish "flexicurity" model relies strongly on public spiritedness. A lack of public spiritedness would raise moral hazard, which would hinder the implementation of an efficient programme of public employment insurance (Algan and Cahuc 2005). On the other hand, informal institutions can impair the effectiveness of formal institutions. If for example the population has strong feelings about "fairness" in the sense of a low wage spread, then wage reductions as a result of decentralising wage negotiations will be difficult to attain (Bewley 2004).

\* Wolfgang Ochel is senior researcher at the Ifo Institute for Economic Research at the University of Munich and CESifo Research Fellow. Gratitude is expressed to the participants of the IAB/IWH Workshop on Labour Market Institutions and Employment in Halle on November 14 and 15, 2005 and especially to Werner Eichhorst for valuable comments.

The nature of the legal system affects, to a large extent, the character of institutional systems. This influence is quite distinct, depending on whether the legal system is based on English common law or whether it is based on civil (or statutory) law. Common law is characterised by the importance of decision making by juries, by independent judges, emphasis on judicial discretion as opposed to the dominance of codified law. The system of common law evolved originally in England and was transplanted to English-speaking countries. Civil law is characterised by a less independent judiciary and gives a greater role to codified substantive and procedural rules. It evolved out of Roman law and has been incorporated into the civil codes of France and Germany and taken over by many countries on the European continent and by Japan (Botero et al. 2003, 7-9).

Following Botero et al. (2003), formal labour market institutions may be assigned to the categories of employment law, of industrial and collective relations law, and social security law. Employment laws govern the individual employment relationship, including the nature of labour contracts, the terms of the contract and the termination of the employment relationship. Laws on industrial relations and collective relations regulate the process of wage bargaining and the adoption and enforcement of collective agreements; they also form the legal basis for trade unions and they lay down the framework for workers' or employers' industrial action. Social security laws govern the social response to individual needs. They deal with old age, disability, illness and unemployment (Table 1). In addition to the three areas of regulation just mentioned, in the literature on this subject active labour market policy (in addition to the already mentioned passive measures) and taxation of labour income are also reckoned among labour market institutions affecting the development of labour markets (Nickell et al. 2005; Checchi and Lucifora 2002, 374). Occasionally regulations of product markets, barriers to labour mobility, the institutional system relevant to private households and the system of vocational and further training are also included amongst labour market institutions (Schütz et al. 1998).

**Table 1**  
**Statutory labour regulations**

<b>Employment laws</b>	
<i>Alternative employment contracts</i>	
<ul style="list-style-type: none"> <li>- Part-time contracts</li> <li>- Fixed-term contracts</li> <li>- Family member contracts</li> </ul>	
<i>Conditions of employment</i>	
<ul style="list-style-type: none"> <li>- Flexibility of working time requirements</li> <li>- Mandatory payment for non-working days</li> <li>- Minimum wage legislation</li> </ul>	
<i>Job security</i>	
<ul style="list-style-type: none"> <li>- Grounds for dismissal</li> <li>- Procedures for dismissal</li> <li>- Notice period</li> <li>- Severance payment</li> <li>- Constitutional principles covering protections against dismissal</li> </ul>	
<b>Industrial (collective) relations law</b>	
<i>Collective bargaining</i>	
<ul style="list-style-type: none"> <li>- Duty to bargain with unions</li> <li>- Extension laws</li> <li>- Closed shops</li> </ul>	
<i>Workers' participation in management</i>	
<ul style="list-style-type: none"> <li>- Mandatory appointment of workers to the board of directors</li> <li>- Workers council by law</li> </ul>	
<i>Collective disputes</i>	
<ul style="list-style-type: none"> <li>- Legal strikes</li> <li>- Procedural restrictions to strikes</li> <li>- Employer defences</li> <li>- Compulsory arbitration</li> <li>- Constitutional protection of the right to strike</li> </ul>	
<b>Social security laws</b>	
<i>Old age, disability and death benefits</i>	
<ul style="list-style-type: none"> <li>- Required time of contributions to access a benefit</li> <li>- Contributions as a percentage of the worker's monthly salary</li> <li>- Replacement rate</li> </ul>	
<i>Sickness and health benefits</i>	
<i>Unemployment benefits</i>	
} similar methodology for sub-indices	

Source: Botero et al. 2003.

**Data sets**

For a long time labour market effects of institutions received scant attention. As a result, the task of collecting information on labour market institutions was neglected. Before the 1990s, very little work on the measurement of labour market institutions was carried out. Since then work in this area has been more extensive and efforts have been made to collect data that are internationally comparable. Considerable improvement in the quality of measurement has been made. International organisations, and in particular the OECD, have played an important role in this work. Theoretical and empirical research in the social sciences has also contributed to progress in this area.

In Table 2 the principal sources of internationally comparable data on formal labour market institutions provided by international organisations are listed. The following data sets are of particular interest:

- Employment protection is captured by the OECD Employment Outlook 1999 and 2004 for



the late 1980s, the late 1990s and 2003. The employment protection of regular workers against individual dismissal, the specific requirements for collective dismissals and the regulation of temporary forms of employment are summarised by means of 22 (in some cases 18) single indicators.

- The characteristics of wage-setting institutions (trade union density, collective bargaining coverage, the centralisation and coordination of wage bargaining) are analysed in the OECD Employment Outlook 1994, 1997 and 2004.
- Social security laws are described in the European Commission's MISSOC database and by Social Security Programs Throughout the World. The effects of social benefits on incomes of working-age individuals and their families are analysed in the OECD study Benefits and Wages (2004a). Information on active labour market policies is included in the OECD Employment Outlook.
- The best information on labour taxes is provided by the OECD in Taxing Wages (2005).

Comparable information on labour market institutions are also supplied by the reports and databases of the Fraser Institute, the Heritage Foundation, Incomes Data Services, the International Institute for Management Development, Lausanne, Watson Wyatt Data Services and the World Economic Forum.

Both the quantity and the quality of information on formal labour market institutions have improved in recent years. But there are also lacunae in available data that need to be closed. With respect to some institutional areas, such as court decisions, there is only scarce information. Information on the degree of implementation of institutional arrangements is lack-

ing, as is information on the number of persons that are affected by certain arrangements. Then too, the concepts on which the collection of information is based must be further developed. And finally, the methods of measurement must be improved.

**Measurement**

As a rule, the measurement of labour market institutions is carried out with a certain problem in mind.

**Table 2**  
**Sources of data on formal labour market institutions for international comparisons**

<p><b>General</b></p> <ul style="list-style-type: none"> <li>- OECD Employment Outlook, various issues</li> <li>- OECD Economic Department Working Papers</li> <li>- European Commission, Employment in Europe, various issues</li> <li>- European Industrial Relations Observatory (EIRO)</li> <li>- EUROSTAT, New Cronos database</li> <li>- ILO databases</li> <li>- CESifo DICE database for institutional comparisons</li> </ul>
<p><b>Employment laws</b></p> <p><i>Employment contracts and conditions of employment</i></p> <ul style="list-style-type: none"> <li>- EIRO, Comparative studies and EMIRE</li> <li>- OECD Employment Outlook, various issues</li> </ul> <p><i>Job security</i></p> <ul style="list-style-type: none"> <li>- OECD, Employment Outlook 1999, ch. 2 and 2004, ch. 2</li> <li>- World Bank, Doing business</li> </ul>
<p><b>Industrial (collective) relations law</b></p> <p><i>Collective bargaining</i></p> <ul style="list-style-type: none"> <li>- OECD, Employment Outlook 1997, ch. 3 and 2004, ch. 3</li> <li>- EIRO, Comparative studies and EMIRE</li> <li>- Blanpain E., ed., International Encyclopedia for Labour Law and Industrial Relations</li> <li>- European Industrial Relations Review</li> </ul> <p><i>Workers' participation in management and collective disputes</i></p> <ul style="list-style-type: none"> <li>- Blanpain E., ed., International Encyclopaedia for Labour Law and Industrial Relations</li> <li>- EIRO, Comparative studies and EMIRE</li> <li>- European Industrial Relations Review</li> </ul>
<p><b>Social security laws and active labour market policies</b></p> <ul style="list-style-type: none"> <li>- European Commission, Mutual Information System on Social Protection in the EU Members States and the EEA (MISSOC)</li> <li>- Social Security Programs Throughout the World</li> <li>- OECD, Benefit and Wages, various issues</li> <li>- International Social Survey Programme</li> <li>- OECD Employment Outlook, various issues</li> </ul>
<p><b>Labour taxes</b></p> <ul style="list-style-type: none"> <li>- OECD, Taxing Wages, various issues</li> </ul>
<p>Comparable information on labour market institutions are also supplied by the reports and databases of the Fraser Institute, the Heritage Foundation, Incomes Data Services, the International Institute for Management Development, Lausanne, Watson Wyatt Data Services and the World Economic Forum.</p> <p>Individual researchers have made important contributions on the concept and measurement of labour market institutions. See the references in: Boeri et al. (2001); Kenworthy (2001b); OECD, Benefits and Wages 2004a and OECD, Employment Outlook 2004, ch.2 and 3.</p>

Source: Own compilation.

The first step is to define the institution. Then one proceeds to make the concept on which the collection of information about the institution is based susceptible to measurement. For this purpose, the separate dimensions of the concept must be worked out in greater detail. In the case of employment protection these would be the period of notice required, the amount of severance pay, the definition of unfair dismissal etc.

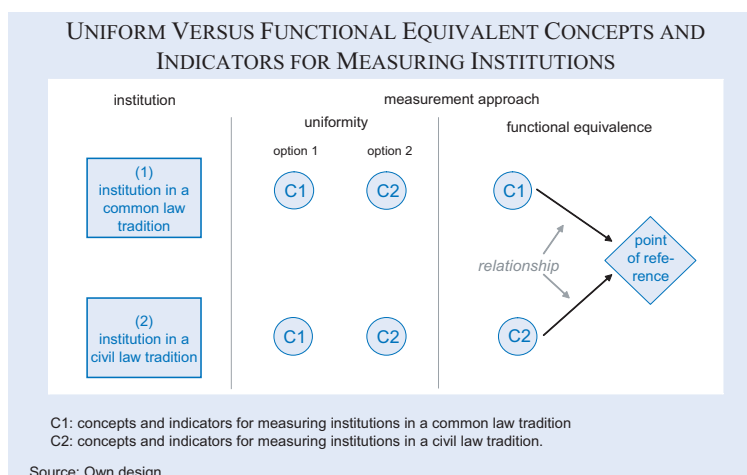
The method of capturing formal institutions consists in the summing-up, interpretation and assessment of laws, ordinances and court decisions by experts. An example of what is meant by summing-up and interpretation is provided by the OECD’s description of employment protection regulations (OECD 2004b, background material for chapter 2). As a rule, assessments are made by assigning scores. Since labour market institutions are typically multidimensional, the task of reducing them to quantitative indices is not simple. Scores may be assigned along a metric scale (e.g. the strictness of employment protection along a scale of one to six, with higher scores representing stricter regulation; OECD 2004b, Annex 2. A1). Or the scores may be based on rank (OECD 1999, tables 2.2–2.5). Internationally comparable assessments may be carried out centrally, or decentrally at the level of individual countries.

In some cases, individual indicators are aggregated to form a composite indicator. Owing to their ability to integrate large amounts of information into an easily understood result, such composite indicators are useful. In constructing composite indicators the relevant indicators are standardised in order to allow comparisons. The indicators are weighted according to their significance and then aggregated (Freudenberg 2003). For instance, in constructing a summary measure for the strictness of employment protection, the OECD started from 18 indicators, which were converted into cardinal scores ranging from 0 to 6. These indicators were then aggregated in a four-step procedure based in this case on an arbitrary weighting scheme. A variety of difficulties can arise when constructing a composite indicator. Outcomes and country rankings may depend largely on the approach selected. For this reason, sensitivity tests should be conducted to analyse the impact of using different stan-

dardisation techniques, changing weights, etc. on the results of the composite indicator.

The measurement of labour market institutions should be as objective, reliable and valid as possible. Objectivity expresses the extent to which the results of measurement are independent of the person that uses the instrument. Reliability is an indication of whether the results of measurement can be duplicated. Validity refers to the extent to which theoretical concepts are captured by the indicators (Dieckmann 1998, 216-227), a criterion that is particularly important. We therefore discuss the problems that can occur with the measurement of labour market institutions with regard to this criterion.

Problems of validity may arise in collecting internationally comparable data on labour market institutions, whenever institutions have evolved in different contexts. An example of this is given by institutions that have evolved in a society with a common law tradition as opposed to the civil law tradition of continental European countries. In such a case a uniform concept and similar indicators do not necessarily adequately reflect the institutions under study. For example, the strictness of employment protection in the Anglo-Saxon countries cannot be registered with the use of indicators that are primarily geared to codified laws. And vice versa, it would not be suitable to examine the dismissal protection regulations of continental European countries using indicators that are primarily based on legal precedents (court decisions). One approach to overcoming this problem is to replace the identity of concepts and indicators by the functional equivalence of concepts and indicators (Kenworthy and Kittel 2003, 22). “Functional equivalence refers to the requirement that concepts (in one setting, W.O.) should be related to concepts in other settings in more or less the



same way” (van Deth 1998, 6). The comparability of measuring concepts can be derived only from their relationship to a common point of reference (Figure). Measurement concepts are equivalent to the degree to which “[the] results provided by [them, W.O.] reliably describe with (nearly) the same validity a particular phenomena in different social systems” (Przeworski and Teune 1970, 108).

Furthermore the validity may be impaired by the method of data collection employed. If the assessment of institutional arrangements in different countries is carried out centrally, e.g. by a supranational organisation, then the problem can arise that national surveys, reports and other source materials on which the assessment is based have been prepared for purposes which differ from the assessment and which furthermore differ from one country to another. As a consequence, the information in the individual countries may be different with respect to its type, content and breakdown. What is more, the availability of information may be very different from one country to another. This may refer to the content of regulatory measures, to the degree to which they are actually implemented and to the number of persons that are affected by them. All of these factors make it difficult to assess national sources. In many cases, organisations conducting central assessments are forced to work with available information on a lowest common denominator basis. Such problems do not arise when specific international surveys are used as a basis of the assessment.

If the assessment of institutional conditions is carried out decentrally by experts in each country, then the assessment is facilitated, since as a rule, local experts have more intimate knowledge of conditions in the country than foreign experts. But different kinds of problems arise in this case. In making international comparisons it is difficult to ensure that national experts employ the same standard in assessing institutional arrangements and that when they assign scores they take into account the relative position of a country in relation to other countries. These difficulties are exemplified by the assessment of hiring and firing practices in different countries presented in the Global Competitiveness Report 2004–2005 of the World Economic Forum (2004). The Forum surveys business leaders with respect to their assessment of hiring and dismissal practices in their own countries (World Economic Forum 2004, 599); the questions are couched in terms of a scoring on a scale from one (hiring and firing of workers is impeded by regulations) to seven (... is flexibly determined by employers). With a

score of 2.2, Germany occupies place 102 among 104 countries, just ahead of France (place 103), but behind Portugal, Spain, Greece and Turkey. At the same time, the World Bank (2005) in its Doing Business Report for 2006 and the OECD (2004b, 117, column 13) both consider Germany’s hiring and firing regulations to be less restrictive than those of the countries just mentioned. The difference in ranking may be due to the assessment being based on different conceptual approaches. But one cannot exclude the possibility that the German experts in the World Economic Forum approached the task of assessing Germany’s labour market flexibility in a more “pessimistic” frame of mind than their foreign counterparts with respect to their own countries. In such a case one might attempt to make the assessment standard more comparable between the countries involved by engaging the experts in an organised exchange of views.

And finally, the method employed in forming composite indicators may give rise to problems of validity. Indicators which are aggregated to a composite indicator have to be weighted. They may be given equal weights or differing weights reflecting their significance. The relative economic impact of the base indicators can be determined by economic theory or by empirical analysis (e.g. regression analysis, principal component analysis or factor analysis) examining the interrelationship among these indicators. Weights can also be set based on correlation coefficients between indicators and a dependent variable such as strictness of employment protection (Freudenberg 2003, 12). Indicators can also be weighted by experts who understand the data and are familiar with the theoretical context. But often transparency is not present. In many composite indicators all base indicators are given the same weight largely for reasons of simplicity. Inappropriate weighting may result in misleading composite indicators.

### Employment protection

In addition to the solution of the measurement problems, the analysis of labour market institutions presupposes that suitable analytical concepts be developed for the individual regulatory areas. In the following the most important concepts for four regulatory areas will be introduced and discussed. We begin with the area of employment protection.

According to the OECD (1999, 50), the term “employment protection” refers both to regulations con-

cerning hiring as well as firing. In the first instance, the relevant regulations concern the conditions under which fixed-term contracts may be concluded, which offer the possibility of circumventing the provisions of protection against dismissal within a regular employment relationship. Regulations with respect to dismissal concern both the individual termination of a regular employment relationship and collective dismissals. The protection of regularly employed workers against dismissal represents a restriction on employers, who are no longer free to give notice to their employees without justification. This restriction has been attained through two types of sanctions: the obligation to continue the employment relationship despite notice having been given or severance pay. The prior condition for the general protection against wrongful dismissal to be effective is that an employment relationship should in fact exist, i.e. that someone is in a position of dependent gainful employment. And finally, there are certain conditions that must be fulfilled if collective dismissal is to be legally justified.

Capturing employment protection is difficult. The arrangements that exist as a result of constitutional provisions, legal measures or collective agreements are complex and the documentation of their implementation is incomplete. The complexity becomes apparent when for example the OECD employs not

less than eight indicators of protection against dismissal of employees with regular employment contracts: notification procedure; delay involved before notice can be given; length of notice period; severance pay; definition of unfair dismissal; length of trial period; compensation following unfair dismissal; and possibilities of obtaining reinstatement after unfair dismissal (Table 3). In order to identify the provisions applicable in this area it is necessary to analyse very carefully the laws, ordinances and wage agreements. But this is only the first step; one must also take into the account how these provisions are implemented and enforced. And this is up to courts, arbitration boards and the public administration in general. Courts of law, for example, interpret how the law is to be applied, decide on the reinstatement of employees in the event of wrongful dismissal, and determine the amount of severance pay, etc. Furthermore, it is of interest to know what proportion of employees take legal action in a court of law to make good their right to seek protection against wrongful dismissal; it is equally interesting to know how often such legal action is successful. There is a similar need for information about the decisions of arbitration boards and the public administration. Administrative records represent an important source of information with respect to the implementation and enforcement of employment protection (Bertola et al. 1999 and 2000).

**Table 3**  
**Employment protection legislation summary indicator at four successive levels of aggregation<sup>a)</sup> and weighting scheme**

Level 4 Scale 0-6	Level 3 Scale 0-6	Level 2 Scale 0-6	Level 1 Scale 0-6	
Overall summary indicator	Regular contracts (5/12)	Procedural inconveniences (1/3)	1. Notification procedures (1/2) 2. Delay to start a notice (1/2)	
		Notice and severance pay for no-fault individual dismissals (1/3)	3. Notice period after 9 months (1/7) 4 years (1/7) 20 years (1/7) 4. Severance pay after 9 months (4/21) 4 years (4/21) 20 years (4/21)	
		Difficulty of dismissal (1/3)	5. Definition of unfair dismissal (1/4) 6. Trial period (1/4) 7. Compensation (1/4) 8. Reinstatement (1/4)	
	Temporary contracts (5/12)	Fixed-term contracts (1/2)	9. Valid cases for use of fixed-term contracts (1/2) 10. Maximum number of successive contracts (1/4) 11. Maximum cumulated duration (1/4)	
		Temporary work agency (TWA) employment (1/2)	12. Types of work for which TWA employment is legal (1/2) 13. Restrictions on number of renewals (1/4) 14. Maximum cumulated duration (1/4)	
	Collective dismissals (2/12)		15. Definition of collective dismissal (1/4) 16. Additional notification requirements (1/4) 17. Additional delays involved (1/4) 18. Other special costs to employers (1/4)	
	<sup>a)</sup> Version 2.			

Source: OECD Employment Outlook 2004, p. 106.

If capturing employment protection in one country is a problem, then obtaining data for international comparisons is all the more difficult, since one must make sure that the information used should be internationally comparable. If one goes beyond cross-section comparisons and attempts a panel analysis, then the concept employed in capturing employment protection must be adjusted to take into account its evolution in the course of time. Basic changes in the regulatory framework must be taken into account as must the emergence of new forms of employment relationships such as fixed-term contracts. As far as possible, one must ensure that concepts are functionally equivalent.

One approach to quantifying the strictness of employment protection in inter-country comparisons is to use surveys. The survey results form the basis for rankings. Such surveys were carried out for the first time in 1985. The International Association of Employers commissioned surveys in 14 countries designed to assess the severity of rules restraining the termination of employment contracts. In the same year, the Commission of the European Union conducted a survey of entrepreneurs in 9 EU countries. In this survey the respondents were asked to assess the employment effect of shorter periods of notice of dismissal, of simpler legal procedures, and of a reduction in redundancy payments (Emerson 1988, reviews the results of these surveys). Bertola (1990) based his rankings of ten industrial countries on the information obtained from these surveys. At the present time, organisations such as Watson Wyatt Data Services, Incomes Data Services and the World Economic Forum carry out surveys.

Whilst the surveys mentioned above request information about the general assessment of the strictness of employment protection, the OECD's work in this area has been based on a number of indicators. Taking Lazear (1990) as a point of departure, who only considered two obstacles to firing workers, Grubb and Wells (1993) and the OECD Job Study (1994b) considered eight indicators referring to obstacles to dismissal of employees with regular contracts (indicators one to eight in Table 3). They also consider the possibilities of circumventing general protection of employment by means

of fixed-term contracts and temporary work agency employment. Regulatory efforts in these two areas are represented by a further six indicators (indicators nine to fourteen in Table 3). In the OECD's Employment Outlook 1999 and 2004 these studies have been broadened by the inclusion of indicators bearing on collective dismissal (indicators fifteen to eighteen in Table 3). The descriptions of these 18 indicators are based on a variety of national sources as well as multi-country surveys by Watson Wyatt Data Services, Incomes Data Services and the European Commission. OECD governments provided additional information based on a request for information from the OECD Secretariat (OECD 1999, 90).

In order to allow for meaningful comparisons, a four-step procedure has been developed for constructing cardinal summary indicators of strictness of employment protection. The 18 indicators are initially expressed in units of time (e.g. months of notice), as a cardinal number (e.g. maximum number of successive fixed-term contracts allowed), or as a score on an ordinal scale (0 to 2, 3, 4 or simply yes/no). These first-level measures are accounted for in comparable units and then converted into cardinal scores ranging from 0 to 6. This scoring algorithm is somewhat arbitrary (OECD 1999, Table 2.B.1 and OECD 2004b, Table 2.A.1.1). The three remaining steps consist in forming successive weighted averages, thus constructing three sets of summary indicators that correspond to successively more aggregated measures of strictness of employment protection (OECD 1999, Annex 2 B; OECD 2004b, Annex 2.A.1; and Table 3).

The OECD summary indicators of the strictness of employment protection rank the United States, Canada, the United Kingdom, Ireland and New Zealand as the OECD member countries providing in 2003 the least

**Table 4**  
**OECD summary indicators of the strictness of employment protection legislation, 2003<sup>a)</sup>**

Country	Score <sup>b)</sup>	Country	Score <sup>b)</sup>	Country	Score <sup>b)</sup>
United States	0.7	Czech Republic	1.9	Norway	2.6
Canada	1.1	Korea	2.0	Sweden	2.6
United Kingdom	1.1	Slovak Republic	2.0	France	2.9
Ireland	1.3	Finland	2.1	Greece	2.9
New Zealand	1.3	Poland	2.1	Spain	3.1
Austria	1.5	Austria	2.2	Mexico	3.2
Switzerland	1.6	Netherlands	2.3	Portugal	3.5
Hungary	1.7	Italy	2.4	Turkey	3.5
Denmark	1.8	Belgium	2.5		
Japan	1.8	Germany	2.5		

<sup>a)</sup> Summary indicator for regular and temporary employment and collective dismissals. – <sup>b)</sup> Higher scores represent stricter regulation.

Source: OECD (2004b, 117).



employment protection. The results of the OECD survey indicate that the strictest protection against dismissal is to be found in three southern European countries: Greece, Spain and Portugal and in the threshold countries Mexico and Turkey (Table 4). For the 28 countries shown in Table 4, the country rankings proposed by the OECD differ considerably from the rankings of the World Economic Forum (2004, 599). Spearman's coefficient of rank correlation is only 0.58.

The indicator of the strictness of employment protection developed by the OECD is in all likelihood the best indicator that is available at the moment for the purpose of making international comparisons in this area. Important areas of regulation are taken into account. The choice of 18 indicators goes far to take adequately into account the complexity of the problem. Nevertheless, the OECD's approach does have some weaknesses:

- The OECD focuses on laws and ordinances bearing on protection against wrongful dismissal, but devotes little attention to other areas such as the system of social security which also may provide protection against loss of employment. One such mechanism is the system of experience rating in the United States where an employer's social security contribution depends in part on the firm's lay-off activity. Then too, the interaction of the protection against dismissal with other labour market institutions must be taken into account if the actual level of protection is to be determined. As Belot and van Ours (2000) have shown, such interactions may reinforce or undermine the level of protection.
- The OECD's measure of employment protection is mainly based on legislative provisions. Protection against dismissal that is a part of wage agreements or of individual employment contracts (e.g. provisions for severance pay) is neglected.
- Similarly, the question to what extent the employment protection legislation is actually enforced receives too little attention. Up till now there has not been an adequate response to Bertola et al.'s (1999) plea for the enforcement of employment protection to be taken into account. The implementation of regulatory measures that are based on legal dispositions is primarily in the hands of labour tribunals. They interpret the law and hand down decisions on the cases brought before them. The stringency of the employment protection actually afforded to workers depends to a great extent on these decisions. The importance of labour tribunals, however, varies greatly from one country to another. According to a survey conducted by the ILO in Spain in 1995, five employees out of a thousand instituted proceedings in a labour tribunal, whilst in Ireland only one out of a thousand took such action. In Spain 72 percent of those who took legal action received a favourable verdict, whereas in Ireland the employees won in only 16 percent of the cases (Bertola et al. 1999, 23). Other disputes are resolved by arbitration boards. It is difficult to collect systematic information on judicial and other resolution of labour disputes (e.g. on the number of cases in litigation, how long they are pending and how they are resolved) and work in this area has only just begun.
- The OECD provides no information on the proportion of employees that are covered by employment protection. It thus does not take into account that legal provisions, wage agreements, court decisions etc. exist which preclude giving regular notice of dismissal to certain clearly defined categories of employees (e.g. older employees, or those who have worked in the production unit for a certain period). On the other hand, it does not take into account that the application of employment protection may depend on the production unit being larger than a minimum size and/or that there may be provisions requiring a waiting period; persons economically active in a production unit that have the formal legal status of self-employed (e.g. a subcontractor) but are deemed to be dependent employees or workers in the informal sector may not be covered by the employment protection provisions either (Rebhahn 2003, 190-194).
- Converting the first-level indicators of employment protection legislation into cardinal scores and the assignment of weights is somewhat arbitrary (Addison and Teixeira 2001, 10-14). "The assignment of scores and weights adds a subjective dimension to the EPL strictness scores that is additional to the judgements already embodied in the... descriptive indicators" (OECD 1999, 117). The extent to which the OECD has analysed the interrelationship among the first-level indicators empirically is not clear.
- The OECD indicator for employment protection only covers the late 1980s, the late 1990s and 2003. In order to be able to carry out panel analyses, it would be desirable if the OECD provided longer and more complete time series.
- Theoretical studies emphasise the analogy between employment protection regulation and a tax borne by the employer on employment adjustment. The cost implications of the various regulatory provisions for employees are not mea-

sured by the OECD. These costs include severance payments, costs of litigation, and costs arising from legally proscribed periods of notice, social plans, and continued payment of remuneration for employees enjoying protection. Furthermore, there are costs that are borne by society in general such as unemployment benefits (Jahn 2004, 11). Information on the costs involved in hiring and firing for businesses are, however, provided by other organisations such as the World Bank Group (2005).

### Wage setting institutions

Collective bargaining needs to be seen against the background of wage setting institutions. The extent of trade union membership and the recognition of unions as a bargaining agent are to a substantial degree determined by regulations. Union bargaining power is normally measured by trade union density and collective bargaining coverage. In addition to these two indicators, the characteristics of the bargaining process play an important role in the evolution of wages. Centralisation and coordination of wage bargaining are considered to be the most important ones.

Trade union density is defined as the ratio of union members to employed wage and salary workers. Gross density refers to all union members, including unemployed and retired members; net density refers only to employed union members. To measure union bargaining power, net density is more appropriate (Ebbinghaus and Visser 2000). In Belgium, Denmark, Finland and Sweden union density is much higher than in other countries. This is due to the so-called Ghent system, whereby unemployment benefits are administered by union-affiliated institutions. Taken in isolation union density is not an adequate measure of union bargaining power. It does not capture the countervailing power of employer associations, the degree of competition in the relevant product markets and the coverage of collective agreements (Flanagan 1999).

Collective bargaining coverage is defined as the proportion of employees (or production units) whose remuneration is regulated by collective wage agreements. The wages agreed upon – the so-called union scales – represent the minimum remunerations. As a rule, the application of a wage agreement is not limited to union members. It has become common practice for employers to apply the terms of collective

contracts to their non-union work force as well. Otherwise, they might be indirectly promoting unionisation. In addition to voluntary extensions, collective agreements can be generally binding within an industrial sector by administrative extensions, covering all employers who are not members of its signatory parties. The scope of collective agreements can also be reduced. Contractual “opt-out clauses” give management and work councils in individual plants authority to make wage agreements. In addition to these legally defined ways, management and work councils can agree upon deviations from currently valid wage agreements without the involvement of the collective bargaining parties or management can breach the wage agreement unilaterally (Ochel 2005). Figures on collective bargaining coverage are usually hard to obtain. For some countries, survey data exist. In other countries, data on coverage are provided by bargaining parties. In several countries, no kind of systematic data collection is undertaken. In this case, experts have to estimate coverage. Information on illegal practices involving deviations from collective agreements is not available (EIRO 2002; OECD 2004b).

The degree of centralisation of wage bargaining refers to the level at which wages are bargained or set. Two elements must be considered in determining the degree of centralisation. The vertical dimension has to do with the aggregation of economic activities: wages can be defined at the company/plant level (decentral), branch/industry level (sectoral) and at the level of the entire economy (central). The horizontal dimension refers to whether workers in different types of jobs (white-collar and blue-collar jobs, different crafts or occupations etc.) bargain jointly or separately. The classification of countries with respect to their bargaining level is complicated by the fact that two or more levels may coexist and are mutually exclusive. Or bargaining may occur at multiple levels, in which case the results of negotiations at the higher levels have a determining effect on the agreement at lower levels (Traxler et al. 2001, 112).

In order to determine the degree of centralisation of wage bargaining it is necessary to obtain information with respect to the number of employees whose remuneration is set at the different levels of negotiation. One approach is only to take into account employees who receive wages or salaries corresponding to union scales. Another – more appropriate and more sensible – approach is to include employees who receive remuneration above union scales (wage

drift) or whose pay is not covered by wage agreements. In such cases effective wages are negotiated individually at the plant level.

The first studies dealing with the degree of centralisation of wage negotiations were carried out by Cameron (1984), Calmfors and Driffill (1988) and others. In the 1990s the OECD estimated the degree of centralisation of wage bargaining in member countries for the years 1980, 1990 and 1994. The OECD’s approach was to assign country scores of 1 (company level), 2 (sector level) and 3 (economy-wide bargaining). Intermediate scores were assigned in cases in which bargaining occurred at more than one level. Ochel (2000) carried the OECD studies a step further by analysing the period from 1960 to 1999. In 2004 the OECD presented a fivefold classification scheme for the wage negotiation systems of 25 countries, covering the period from 1970 to 2000 (OECD 2004b).

In contrast to the centralised approach of the OECD, Traxler et al. pursue a decentral approach in which labour relations experts from 20 countries assessed the wage bargaining system of their country. The assessments were then compared with the existing literature and discrepancies resolved via discussion with the specialists. Apart from the vertical dimension of centralisation (3 levels and mixtures of levels), the horizontal dimension has been also included. Bargaining at a given level may occur for all groups of workers jointly or may be specific to a group. In all there are twelve categories (Table 5). The number of employees covered by each of the different levels is the most relevant criterion for classification. In 1997-98 the United States, Canada and New Zealand had the most decentralised bargaining

system (with a score of 1) while Ireland (with a score of 12) and Finland (11) had the most centralised systems (Traxler et al. 2001, 114).

Whereas the OECD and Traxler et al. (2001) focus on the actual level at which bargaining takes place, Iversen (1999) and Golden et al. (2002) focus on the structural characteristics of the wage-bargaining process. (A survey in greater depth of these approaches can be found in Kenworthy 2001b.) Iversen seeks to identify the locus of bargaining authority. His centralisation index combines a measure of organisations and the share of unionised workers at each of the three main bargaining levels. The capacity of bargaining agents to implement their agreements is taken into account as well. Enforceable agreements presuppose that bargaining agents control most strike and lockout funds and can impose fines for non-compliance so that low-level bargainers cannot circumvent central or industry level wage agreements. By including the enforceability of bargaining agreements, aspects of wage coordination are combined with elements of centralisation of wage bargaining (Iversen 1999, 83-86).

In much the same way as Iversen, the centralisation index of Golden et al. (2002) is a measure of the centralising activities of confederations, rather than of the degree of wage centralisation itself. Their first indicator is an index of involvement in wage bargaining by peak-level union and employer confederations (with scores ranging from 1 to 11). The second is an index of government involvement in the wage setting process (with scores ranging from 1 to 15). The third is a summary index of the overall degree of wage centralisation (with scores ranging from 1 to 4). The Golden-Wallerstein-Lange index is not a pure centralisation index but includes elements of wage coordination. It is the only index taking government-imposed centralisation and coordination explicitly into account. Golden et al. make own assessments of the centralisation of wage bargaining in individual countries and in this respect their approach is similar to the OECD’s and Iversen’s.

Another property of wage bargaining is the degree to which it is subject to coordination. Coordination can be defined as a mechanism to increase the consensus between the participants in the collective bargaining. The degree of coordination reflects the extent to which individual wage settlements are in tune with one another. Or in other words, the extent to which “minor players deliberately follow along with what major players decide” (Kenworthy 2001b, 75).

**Table 5**  
**Centralisation of wage bargaining level in the private sector; scores of the Traxler, Blaschke and Kittel measure<sup>a)</sup>**

Scores	Central	Industry	Company and plant	All groups	Group specific
1			x		x
2			x	x	
3		x	x		x
4		x	x	x	
5		x			x
6		x		x	
7	x	x	x		x
8	x	x	x	x	
9	x	x			x
10	x	x		x	
11	x				x
12	x			x	

<sup>a)</sup> Scores have been reversed so that higher values indicate greater centralisation.

Source: Traxler et al. (2001), 307.

Coordination and centralisation of wage negotiations are not identical. Coordination may occur even though the negotiations are conducted decentrally. Coordination processes are complex. They can only be captured when both horizontal and vertical coordination are taken into account (Moene et al. 1993). Horizontal coordination aims at harmonising wage bargaining across distinct jobs and business activities. Vertical coordination seeks to make the rank and file follow the decisions taken by their representatives at higher levels. Apart from these dimensions of coordination activity it is important that the coordinating activities of the state should not be neglected (Traxler et al. 2001).

Based on a series of studies reflecting research in this area, the OECD has estimated the degree of co-

ordination of wage bargaining in its member countries for the years 1980, 1990 and 1994 (OECD 1994a and 1997). The OECD assigned country scores of 1 (uncoordinated), 2 (medium degree of coordination) and 3 (highly coordinated). Ochel (2000) extended the OECD results to comprehend the period 1960 to 1999. The trinary assessment scheme does not do justice to the complexity of the coordination processes. Neither do these studies provide an explicit rationale for their coding, although many aspects of coordination have been taken into account implicitly.

The different dimensions of coordination are taken into account in OECD (2004b) and in Kenworthy (2001a). Both the dimension of vertical coordination as well as the coordination efforts of the state have their place in the classification of the wage negotia-

**Table 6**  
**Coordination of wage bargaining in OECD countries, 1995-2000**

Country	OECD <sup>a) b)</sup> 2004	Kenworthy <sup>a) c)</sup>	Country	OECD <sup>a) b)</sup> 2004	Kenworthy <sup>a) c)</sup>
Australia	2	2	Korea	1	...
Austria	4	4	Netherlands	4	4
Belgium	(4.5)	(4.5)	New Zealand	1	1
Canada	1	1	Norway	(4.5)	(4.5)
Czech Republic	1	...	Poland	1	...
Denmark	(4)	(3.5)	Portugal	4	...
Finland	5	4	Slovak Republic	2	...
France	2	2	Spain	3	...
Germany	4	4	Sweden	3	3
Hungary	1	...	Switzerland	4	4
Ireland	4	5	United Kingdom	1	1
Italy	4	4	United States	1	1
Japan	4	5			

- ... Data not available.
- a) Figures in brackets are period averages in cases where at least two years differ from the period's modal value.
- b) 1 = Fragmented company/plant bargaining, little or no coordination by upper-level associations.  
 2 = Fragmented industry and company-level bargaining, with little or no pattern-setting.  
 3 = Industry-level bargaining with irregular pattern-setting and moderate coordination among major bargaining actors.  
 4 = a) informal coordination of industry and firm-level bargaining by (multiple) peak associations;  
 b) coordinated bargaining by peak confederations, including government-sponsored negotiations (tripartite agreements, social pacts), or government imposition of wage schedules;  
 c) regular pattern-setting coupled with high union concentration and/or bargaining coordination by large firms;  
 d) government wage arbitration.  
 5 = a) informal coordination of industry-level bargaining by an encompassing union confederation;  
 b) coordinated bargaining by peak confederations or government imposition of a wage schedule /freeze, with a peace obligation.
- c) 1 = Fragmented wage bargaining, confined largely to individual firms or plants.  
 2 = Mixed industry- and firm-level bargaining, with little or no pattern setting and relatively weak elements of government coordination such as setting of basic pay rate or wage indexation.  
 3 = Industry-level bargaining with somewhat irregular and uncertain pattern setting and only moderate union concentration.  
 4 = a) Centralized bargaining by peak confederation(s) or government imposition of a wage schedule/freeze, without a peace obligation (Belgium in most years, and Finland);  
 b) informal centralization of industry- and firm-level bargaining by peak associations (Italy, Netherlands, Norway in some years, Switzerland);  
 c) extensive, regularized pattern setting coupled with a high degree of union concentration (Germany, Austria).  
 5 = a) centralized bargaining by peak confederation(s) or government imposition of a wage schedule/freeze, with a peace obligation (Ireland, Norway in some years);  
 b) extensive, regularized pattern setting and highly synchronized bargaining coupled with coordination of bargaining by influential large firms (Japan).

Sources: OECD 2004 b; Kenworthy (2001 a) and Data set (www.u.arizona.edu/~lkenwor/WageCoorScores.xls).

tion systems. The OECD distinguishes nine different ways of determining wages, which can be summarised in five degrees of coordination (Table 6). In contrast to the OECD, Kenworthy does not measure wage coordination per se. His index with five categories “is instead a hypothesis or prediction about the degree of coordination that is likely to be generated by various wage-setting institutions... Because it focuses on the structural characteristics of the wage-setting process, it is considerably easier to measure than is the degree of coordination the process actually generates” (Kenworthy 2001b, 79-80). Nonetheless, even with such an approach it is necessary to test the hypotheses with respect to the degree of coordination.

Traxler et al. (2001) do not estimate the intensity of coordination directly. Rather, they seek to capture the coordination activities of the organisations actively involved in wage negotiations. Since these activities are qualitatively different from one another, they create a categorical coordination indicator. In their view there are six principal modes of coordination:

- inter-associational coordination by the peaks of unions and employer associations,
- intra-associational coordination by the peaks,
- pattern bargaining,
- state-imposed coordination,
- state-sponsored coordination, and
- uncoordinated bargaining.

Vertical bargaining governability is measured by the ability of higher-level agreements to impose wage moderation on the shop floor. Two types of rules are most conducive to vertical coordination: the legal enforceability of wage agreements and a peace obligation. Vertical bargaining governability is high when both these commitments are guaranteed effectively. Otherwise it is low. Vertical bargaining governability is present only as far as peak-level coordination is concerned. It is insignificant or pointless in the case of

pattern bargaining, state-imposed coordination and uncoordinated bargaining.

Traxler et al. (2001) have analysed the effect of their coordination patterns on labour costs in a panel study. Peak-level coordination backed by high governability as well as pattern bargaining show the lowest increase in labour costs. State-imposed coordination (and uncoordinated bargaining) shows an average performance. Peak-level coordination under low governability leads to the highest increase in labour costs (Table 7).

### Social security and active labour market policy

Social security systems insure workers against the risk of unemployment and redistribute income in favour of the most disadvantaged workers (passive measures of labour market policy). At the same time, social benefits may reduce peoples’ efforts to look for a job and increase the reservation wage, thus exerting upward pressure on wages. These effects may increase the duration of unemployment.

The degree of income maintenance and the financial work incentives are measured by the difference between out-of-work and in-work incomes. The ratio defines the so-called replacement rates. In calculating replacement rates several choices have to be made (Atkinson and Micklewright 1991). If out-of-work benefits are seen as an insurance system, then the numerator would be out of work benefit income and the denominator would be income from work of the person whose labour market status changes. If the living standard out of work as opposed to that in work is of interest, then in addition all other incomes that are independent of work status should be included in the numerator and the denominator. In this case a household concept would be appropriate. In

addition one has to decide on the type of transition: from employment to unemployment or vice versa (or transitions into other labour market states). Finally gross or net replacement rates can be calculated, the latter being the more comprehensive measure (Immervoll and O’Donoghue 2003).

The best known net replacement rates (NRR) are those calculated by the OECD (2004a). They show

**Table 7**  
**Bargaining coordination and wage moderation**

Wage moderation			
Strong		Pattern bargaining	Voluntary peak-level coordination with high bargaining governability
Medium	Uncoordinated bargaining		State-imposed coordination
Weak			Voluntary peak-level coordination with low bargaining governability
	Low	Medium	High
Vertical coordination			

Source: Traxler et al. 2001, p. 247.



the proportion of in-work income that is maintained for someone becoming unemployed. They refer to persons who were previously employed on a full-time basis with earnings at 67, 100 and 150 percent of the average production worker wage. As indicators of net income, they capture the direct effects of all relevant types of taxes and benefits. Furthermore, NRR is calculated for different family types taking into account the household as a whole. Finally, to capture different durations and time profiles of out-of-work benefits, replacement rates are calculated for the initial phase of unemployment as well as for longer periods of joblessness (up to the 60th month of benefit receipt).

Although the OECD net replacement rate is one of the most carefully constructed measuring instruments, it has some shortcomings. It provides no information about the proportion of recipients of income replacement benefits for which a certain replacement rate is relevant. What is more, up till now certain groups of persons have not been included in the calculations. It would, for example, be an advantage if net replacement rates for economically active persons of different age groups would be calculated.

Net replacement rates and duration of entitlement alone are not decisive for the incentive effects of a benefit system. They have to be analyzed together with the eligibility criteria. Eligibility criteria restrict social benefits to people who meet requirements such as independent job search, attending interviews, accepting suitable work etc. They can be made effective by strict legislation and by sanctions. The strictness with which the benefit system is operated is described in MISSOC, OECD (2000, chapter 4) and Grubb (2001). Very little information on the implementation of sanctions is available. Incidences of refusal of benefits are shown in Grubb (2001, Table 2).

Net replacement rates and eligibility criteria represent the financial incentives to work for persons receiving income replacement benefits. In 1999, in 16 OECD countries 18.6 percent of persons of working age fell into this category. Net replacement rates provide no indication of the financial incentives to work for persons who are unemployed, but do not receive income replacement benefits. 21.3 percent of persons of working age fell into this category (OECD 2003, 175).

Besides passive measures, active labour-market policy measures (ALMP) are carried out in all OECD countries. The purpose of ALMPs is to provide ac-

tive assistance to the unemployed, which will improve their chances of obtaining work. The ALMP is registered in terms of expenditure as a share of GDP, of the total budget for labour market policies, etc. Further indicators of ALMPs are the numbers of persons in active labour-market policy measures as well as people who join or leave these programmes. The OECD database on Labour Market Programmes is the most important source. However, it contains no time series that date back to the 1960s and 1970s. Also the departures from active labour market policies are not subdivided according to whether the persons have moved into non-subsidised jobs, into unemployment or into inactivity. This subdivision would allow conclusions to be made on the effectiveness of the ALMP (Eichhorst et al. 2001, 198).

### Taxing labour

The tax burden on labour has substantial policy relevance. As opposed to the burden on capital income it has implications for the distribution of income. Furthermore, it affects the efficiency of the labour market, influencing both participation rates and unemployment rates. One way to calculate the tax burden on labour is the Taxing Wages approach of the OECD (2005). Taxing Wages seeks to determine the combined effect of personal income taxes, social security contributions and family cash benefits on the net incomes of various illustrative family-types and on the labour costs faced by employers. Information is provided on employees at comparable levels of income. The main focus is on the "average production worker". The calculations in Taxing Wages take no account of observed data.

The strength of the Taxing Wages methodology lies in its ability to make international comparisons of tax systems, without being affected by different population structures. However, it is limited by considering a restricted number of household types and a fairly narrow income range. Other limitations are the exclusion of taxes on the goods that workers consume, the exclusion of non-wage income and the limited numbers of tax reliefs covered. As social benefits are not included either, Taxing Wages is not the best source to analyse the incentive for individuals to participate in the labour market (Heady 2003).

The Taxing Wages approach is not the only way in which the taxation of labour income can be assessed. One alternative is to calculate the implicit average

effective tax rate, by estimating the total amount of tax paid on labour earnings in a country and dividing that by an estimate of total wages or labour costs (OECD 2001). Implicit average effective tax rates have the advantage of being based entirely on observed quantities, and thus reflect all the factors that influence the amount of taxes actually paid. Furthermore, they take account of the taxation of all workers. On the other hand, they do reflect a combination of differences in tax systems and differences in the population structure, which makes international comparisons of tax differences more difficult (Heady 2003).

Another alternative is to use a micro-simulation model to calculate labour taxes for a representative sample of a country's population. An example of this is EUROMOD (Sutherland 2001). Micro-simulation models can provide results for each of the households in the database that it uses. In principle, they can produce figures for the taxation of a broad range of incomes, although for high-income individuals information on tax deductions or tax avoidance is rarely sufficient. Similar to Taxing Wages the simulation of taxes paid does not take account of observed data.

### **The need for more information**

The quantity and quality of information on formal labour market institutions have improved considerably in the last fifteen years. Nevertheless more information is still needed.

In the first place, there is a need for more and better information on areas of regulation which up till now have not been the object of systematic data collection. This refers to "new" areas of regulation like "opt-out clauses" that permit area-wide wage agreements to be set aside in favour of agreements at the plant level. Furthermore, it refers to areas that are not well documented. This applies, for example, to the decisions handed down by labour tribunals and arbitration boards with respect to employment protection and to severance pay. Beyond that, increasingly there are institutional areas that need to be the object of systematic data collection on account of their indirect effects. An example of this would be the system of experience rating in the United States and its effect on protection against wrongful dismissal. And finally, there are areas of regulation which have up till now not been subject to systematic data collection such as wage agreements, the internal arrangements

of organisations and the eligibility requirements with respect to transfer payments. Sectoral and regional differences of institutional regulations should also be captured. An improved state of information in this respect could enhance the targeting of labour market politics.

Apart from the inclusion of new areas of regulation, implementation of institutional arrangements should receive more attention. The texts of laws and edicts do not tell us whether they are actually applied or not. One country's strict rules may be paired with lax enforcement, whilst another country's lax provisions may be applied with greater rigour. Information concerning the implementation of rules and regulations is often inadequate. An improvement would require that national administrative entities systematically collect information on their activities and make this information available to international organisations such as the OECD. International surveys could also help to improve the situation with respect to information availability.

Another aspect is the improvement of the methods employed to obtain information on labour market institutions. First of all, the theoretical concepts through which we examine institutions and their activities need to be further developed and applied. Second, the methods of measurement need to be improved. The measurement methods should meet, as far as possible, the criterion of functional equivalence. Whenever decentral assessments of institutional arrangements are undertaken, care should be taken that the national experts base their assessments on homogeneous criteria. Whenever the available data are insufficient, international surveys should be conducted. And, last but not least, in calculating composite indicators, a weighting scheme reflecting the significance of individual indicators should be employed.

Another topic of interest concerns the relevance of individual institutions. How many people make use of the services of an institution and how often do they do this? How many persons are affected by a given regulatory arrangement? Which set of persons is covered by a given indicator? The question of relevance is closely connected to the question of implementation: if a regulation or a regulatory arrangement is not implemented, then it is not relevant.

Furthermore, wherever possible, long time-series on the development of labour market institutions should be provided. These are needed for panel analyses and

to identify the extent to which individual institutions are changeable. Information on the possibilities of changing formal institutions is of great interest. Beyond that, one must attempt to determine the costs which would be incurred in modifying institutional arrangements and in overcoming the resistance to such modifications. Improvement in the availability of information in this area would require considerable preliminary research efforts.

## Summary

The way labour market institutions have come into being and the effects that they have are aspects which have received a good deal of attention, but the collection of information on such institutions has been much neglected. Since the 1990s, however, thanks to the work done at the OECD and other international organisations and thanks to social science research, considerable progress has been made. The concepts underlying the collection of information about labour market institutions have been developed further as has been shown with the examples of employment protection, wage setting institutions, social security and ALMP, and taxation of labour. The methods of measurement have also been improved.

Nonetheless, the need for information is still far from being completely satisfied. The systematic collection of information is still not adequate in many areas of regulatory activity. There is relatively little information on the implementation of regulations. Concepts and methods of measurement must be further developed. In addition, information concerning the relevance of institutional arrangements is needed. And finally, the inertial resistance of institutional arrangement to change requires study.

## References

- Addison, J. T. and P. Teixeira (2001), "The Economics of Employment Protection", *IZA Discussion Paper* 381, Bonn.
- Agell, J. (1999), "On the Benefits from Rigid Labour Markets Norms, Market Failures, and Social Insurance", *The Economic Journal* 109 (February), F 143-F 164.
- Algan, Y. and P. Cahuc (2005), "Civic Attitudes and the Design of Labor Market Institutions: Which Countries can Implement the Danish Flexicurity Model?", September, mimeo.
- Atkinson, A. and J. Micklewright (1991), "Unemployment Compensation and Labor Market Transitions: A Critical Review", *Journal of Economic Literature* 29(4), 1679-1727.
- Belot, M. and J. van Ours (2000), "Does the Recent Success of some OECD Countries in Lowering their Unemployment Rates Lie in the Clever Design of their Labour Market Reforms?", *IZA Discussion Paper* 147, Bonn.
- Bertola, G. (1990), "Job Security, Employment and Wages", *European Economic Review* 34, 851-886.
- Bertola, G., T. Boeri and S. Cazes (1999), *Employment Protection and Labour Market Adjustment in OECD Countries: Evolving Institutions and Variable Enforcement*, Employment and Training Papers 48, ILO, Geneva.
- Bertola, G., T. Boeri and S. Cazes (2000), "Employment Protection in Industrialized Countries: The Case of New Indicators", *International Labour Review* 139(1), 57-72.
- Bewley, T. (2004), "Fairness, Reciprocity, and Wage Rigidity", *IZA Discussion Paper* 1137, Bonn.
- Blau, F. D. and M. Kahn (1999), "Institutions and Laws in the Labor Market", in O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. 3, 1399-1461.
- Boeri, T., A. Brugiavini and L. Calmfors, eds. (2001), *The Role of Unions in the Twenty-first Century*, Oxford.
- Botero, J., S. Djankov, R. La Porta, F. Lopez-de-Silanes and A. Shleifer (2003), "The Regulation of Labor", *NBER Working Paper* 9756.
- Calmfors, L. and J. Driffill (1988), "Bargaining Structure, Corporatism, and Macroeconomic Performance", *Economic Policy* 6, 13-61.
- Cameron, D. R. (1984), "Social Democracy, Corporation, Labour Quiescence, and the Representation of Economic Interest in Advanced Capitalist Society", in J. H. Goldthorpe, ed., *Order and Conflict in Contemporary Capitalism*, Oxford.
- Checchi, D. and C. Lucifora (2002), "Unions and Labour Market Institutions in Europe", *Economic Policy* 17(2), 362-401.
- Diekmann, A. (1998), *Empirische Sozialforschung, Grundlagen, Methoden, Anwendungen*, 4. Auflage, Reinbek bei Hamburg.
- Ebbinghaus, B. and J. Visser (2000), *The Societies of Europe, Trade Unions in Western Europe since 1945*, London.
- Eichhorst, W., S. Profit and E. Thode (2001), *Benchmarking Deutschland: Arbeitsmarkt und Beschäftigung, Bericht der Arbeitsgruppe Benchmarking und der Bertelsmann Stiftung*, Berlin, Heidelberg.
- EIRO (2002), "Collective Bargaining Coverage and Extension Procedures", eiroline, December. [www.eiro.eurofound.ie.int/2002/12/study/tn0212102s.html](http://www.eiro.eurofound.ie.int/2002/12/study/tn0212102s.html).
- Emerson, M. (1988), "Regulation or Deregulation of the Labour Market, Policy Regimes for the Recruitment and Dismissal of Employees in the Industrialised Countries", *European Economic Review* 32, 775-817.
- Flanagan, R. J. (1999), "Macroeconomic Performance and Collective Bargaining", *Journal of Economic Literature* 37, September, 1150-1175.
- Freudenberg, M. (2003), "Composite Indicators of Country Performance: A Critical Assessment", *OECD, DSTI Working Paper* 16, Paris.
- Golden, M., P. Lange and M. Wallerstein (2002), "Union Centralization among Advanced Industrial Societies. An Empirical Study". Dataset available at <http://www.shelley.polisci.ucla.edu/data>. Version dated July 28, 2004.
- Grubb, D. (2001), "Eligibility Criteria for Unemployment Benefits", in OECD, ed., *Labour Market Policies and the Public Employment Service*, Paris, 187-216.
- Grubb, D. and W. Wells (1993), "Employment Regulations and Patterns of Work in OECD Countries", *OECD Economic Studies* 21, Winter, 7-58.
- Heady, C. (2003), "The 'Taxing Wages' Approach to Measuring the Tax Burden of Labour", *CESifo Working Paper* 967.
- Immervoll, H. and C. O'Donoghue (2003), "Employment Transitions in 13 European Countries. Levels, Distributions and Determining Factors of Net Replacement Rates", *CESifo Working Paper* 1091.
- Iversen, T. (1999), *Contested Economic Institutions, The Politics of Macroeconomics and Wage Bargaining in Advanced Democracies*, Cambridge.
- Jahn, E. J. (2004), "Der Kündigungsschutz auf dem Prüfstand", *Arbeitspapier der Konrad-Adenauer-Stiftung e.V.* 138, Sankt Augustin.

- Kenworthy, L. (2001a), "Wage Setting Coordination Scores" (Manuscript; [www.u.arizona.edu/~lkenwor/WageCoorScores.pdf](http://www.u.arizona.edu/~lkenwor/WageCoorScores.pdf)).
- Kenworthy, L. (2001b), "Wage-Setting Measures. A Survey and Assessment", *World Politics* 54, October, 57–98.
- Kenworthy, L. and B. Kittel (2003), *Indicators of Social Dialogue: Concepts and Measurement*, ILO Working Paper No. 5, Geneva.
- Lazear, E. P. (1990), "Job Security Provisions and Employment", *The Quarterly Journal of Economics*, August, 699–726.
- Moene, K. O., M. Wallerstein and M. Hoel (1993), "Bargaining Structure and Economic Performance", in R. J. Flanagan, K. O. Moene and M. Wallerstein, eds., *Trade Union Behavior, Pay-Bargaining, and Economic Performance*, Oxford, 63–131.
- Nickell, S., L. Nunziata and W. Ochel (2005), "Unemployment in the OECD since the 1960s. What Do We Know?", *The Economic Journal* 115 (500), 1–27.
- Ochel, W. (2000), *Collective Bargaining (Centralisation/Co-ordination)*, Ifo Institute, Munich, mimeo.
- Ochel, W. (2005), "Decentralizing Wage Bargaining in Germany – A Way to Increase Employment?", *Labour* 19(1), 91–121.
- OECD (1994a), *Employment Outlook*, Paris.
- OECD (1994b), *The OECD Jobs Study*, Part II, Paris.
- OECD (1997), *Employment Outlook*, Paris.
- OECD (1999), *Employment Outlook*, Paris.
- OECD (2000), *Employment Outlook*, Paris.
- OECD (2001), *Tax Ratios: A Critical Survey*, OECD Tax Policy Studies No. 5, Paris.
- OECD (2003), *Employment Outlook*, Paris.
- OECD (2004a), *Benefit and Wages, OECD Indicators, 2004 Edition*, Paris.
- OECD (2004b), *Employment Outlook*, Paris.
- OECD (2005), *Taxing Wages 2003/2004-2004 Edition*, Paris.
- Przeworski, A. and H. Teune (1970), *The Logic of Comparative Social Inquiry*, New York.
- Rebhahn, R. (2003), "Der Kündigungsschutz der Arbeitnehmer in den Staaten der EU", *ZfA* 34(2), 163–235.
- Sutherland, H. (2001), "EUROMOD, An Integrated European Benefit-Tax Modell", *EUROMOD Working Paper EM 9/01*.
- Schütz, H., S. Speckesser and G. Schmid (1998), "Benchmarking Labour Market Performance and Labour Market Policies: Theoretical Foundations and Applications, Discussion Paper FSI 98–205, Wissenschaftszentrum Berlin für Sozialforschung.
- Traxler, F., S. Blaschke and B. Kittel (2001), *National Labour Relations in Internationalized Markets, A Comparative Study of Institutions, Change, and Performance*, Oxford.
- Van Deth, J.W. (1998), "Equivalence in Comparative Political Research", in J. W. van Deth, ed., *Comparative Politics: The Problem of Equivalence*, London, 1–19.
- Voigt, St. (2002), *Institutionenökonomik*, München.
- World Bank Group (2005), *Doing Business 2006*, Washington D.C.
- World Economic Forum (2004), *The Global Competitiveness Report 2004-2005*, Geneva.

## THE EFFECTIVE TAX BURDEN OF COMPANIES IN EUROPE

MICHAEL OVERESCH\*

Company taxation is commonly accepted as a relevant location factor. In this context the measurement and the international comparison of the effective tax burden indicates differences in the attractiveness of locations. This study compares the effective tax burden of companies based on a measure which reflects the impact of company taxation on decisions and in particular on location choices. The calculations were carried out at the Centre for European Economic Research (ZEW) and are based on an approach which was introduced by Devereux and Griffith (1999). This approach is useful for analysing the impact of taxation on investment decisions. Moreover, profit-shifting strategies can also be integrated. A more detailed study based on the same approach covering regional differences and additional non-EU countries has been done for the IBC International Benchmark Club of BAK Basel Economics.<sup>1</sup>

### Measuring the effective tax burden

Regarding the impact of company taxation on location decisions, a measure of the effective tax burden has to reflect the decision process on investment strategies. These so-called forward-looking approaches calculate the tax burden on a hypothetical investment project of a company taking into account the existing tax rules. In general, so-called backward-looking approaches cannot measure the impact of taxation on decisions (Sørensen 2004, 17–19). Well-known examples for backward-looking measures are the implicit tax rates provided by the EU Communities (2005). They are helpful in analysing distribution effects of taxation but not in the context of looking at company taxation as a location factor. A comparison only based on statutory income tax rates is also insufficient, because this would neglect differences in the determination of tax bases and non-income taxes. Hence, in the context of taxation as a location factor measures of an effective tax burden

should be calculated as a share of an investor's financial target, e.g. the project's net present value. The approach of Devereux and Griffith (1999) used for calculating the effective tax burden of companies in this study fulfils all these requirements. This neo-classical model is based on a commonly accepted framework developed by King and Fullerton (1984). It provides a possibility for taking into account the most relevant provisions of tax regimes in a systematic way. Using this approach, cost of capital, an effective marginal tax rate (EMTR) and an effective average tax rate (EATR) can be computed. The cost of capital and the EMTR are measures for the effective tax burden attributable to marginal investments, whereas the EATR shows the effective tax burden on profitable investments.

Marginal investments display a net present value of zero, i.e. they yield a rate of return on the initially invested capital that is just sufficient in order to compete with the alternative investment. This minimum rate of return before taxes required by an investor is called cost of capital. Thus, in the absence of taxes, the cost of capital equals the real market interest rate. If taxation causes the cost of capital to fall below the real market interest rate, it favours the corporate investment over the alternative investment and vice versa. In this case, taxation exerts an influence on the optimal level of investment activity. Furthermore, the cost of capital can act as an indicator for the competitiveness of a company, since it determines the long-term lower limit of potential prices at which the company can offer its products. While the cost of capital measures the minimum rate of return, the EMTR reflects the percentage difference between the cost of capital, denoted by  $\tilde{p}$  and the post-tax real rate of return, denoted by  $s$ :

$$EMTR = \frac{\tilde{p} - s}{\tilde{p}}$$

The EMTR determines the share of the return on a marginal investment which is cut by taxation. If we focus only on taxation at the corporate level, the real post-tax rate of return equals the real market interest rate  $r$ . In this case, the EMTR and cost of capital contain the same information.<sup>2</sup> Determining the effective tax burden on marginal investments in terms of EMTR facilitates the comparison with other concepts of tax rates like EATR or the statutory profit tax rate.

\* Michael Overesch is researcher at the Centre for European Economic Research (ZEW) in Mannheim.

<sup>1</sup> See IBC Taxation Index 2005 ([www.bakbasel.com](http://www.bakbasel.com)).

<sup>2</sup> However, if personal taxes are considered, cost of capital and the EMTR provide different information.



The EATR reflects the percentage reduction of the net present value of a profitable, inframarginal investment that is caused by taxation. An inframarginal, profitable investment yields a rate of return  $p$  above the cost of capital  $\tilde{p}$ . Detailed technical descriptions of effective tax rates are provided by Devereux and Griffith (1999) or Schreiber, Spengel and Lammersen (2002). When choosing between two or more mutually exclusive profitable investments, a company will favour the alternative that yields the highest post-tax net present value. Location choices for subsidiaries of international corporations represent the most relevant example for this kind of decision. Consequently, the EATR is an important indicator for the attractiveness of a location, whereas the cost of capital indicates the optimal size of an investment.

The following equation describes a particular relationship between the cost of capital, the EMTR and the EATR:

$$EATR = \frac{\tilde{p}}{p} \cdot EMTR + \frac{p - \tilde{p}}{p} \cdot \tau$$

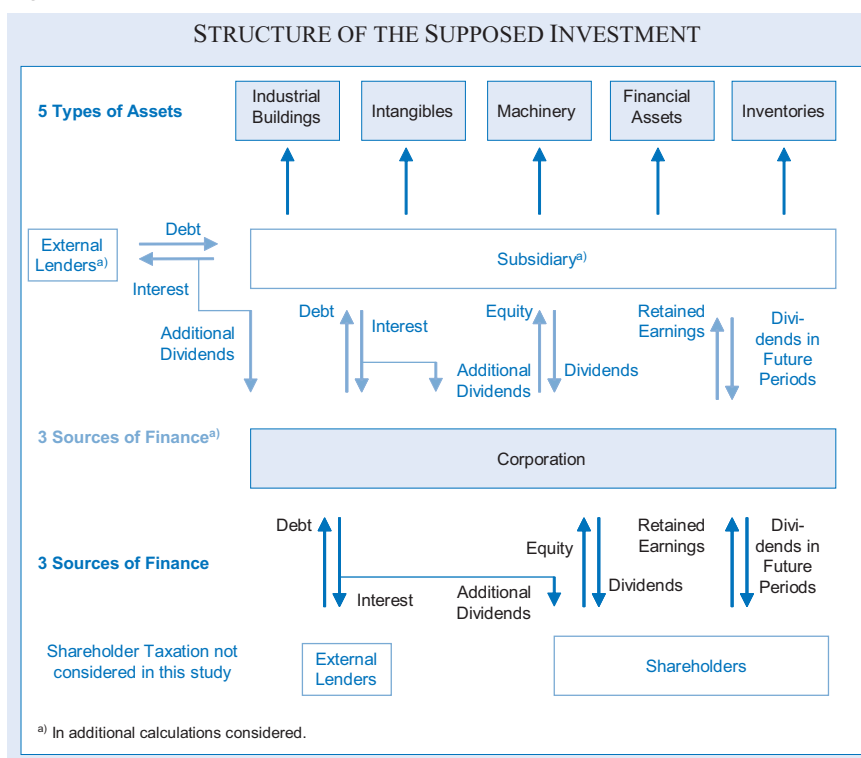
This relationship illustrates the properties of the EATR and helps to identify the impact of the different tax drivers on the effective tax burden. The EATR equals the weighted average of the EMTR and the combined effective statutory corporate income tax rate, denoted by  $\tau$ . The weights are determined by the proportion of the pre-tax return  $p$  that is covered by the cost of capital  $\tilde{p}$  (for the EMTR) and the fraction that is above the cost of capital (for the combined tax rate). Consequently, the EATR equals the EMTR if the assumed rate of return of an additional investment equals the cost of capital. The effective tax rate of an investment does not only depend on the statutory corporate income tax rate, but is also affected by the definition of the tax base – especially by tax depreciation

allowances – and by non-income taxes. However, the more the rate of return exceeds the cost of capital the more the EATR converges against the combined effective statutory corporate income tax rate  $\tau$ . Therefore, if the level of profitability is increased, the treatment of expenses for tax purposes will become less relevant for the determination of the effective tax burden. Since marginal and profitable investments display the same initial cost but different levels of return, non-income taxes cut a lower proportion of the return of a more profitable investment and become less relevant as well. In summary, the statutory income tax rate becomes the dominant factor in determining the effective tax burden of a highly profitable investment.

**Assumptions of the model**

The investment and financial structure of the model is illustrated in Figure 1. The model assumes a company in the manufacturing sector with the legal form of a corporation. This corporation invests in five different assets: industrial buildings, intangibles (patents) bought from third parties, machinery, financial assets, and inventories. The types of assets are weighted equally. The financing policies of the corporation take three different sources of finance into account: new equity capital, retained earnings and debt from external lenders. The sources of finance are also

**Figure 1**



weighted equally. The EATR is calculated by assuming a pre-tax real rate of return of about 20 percent. Note that this study considers taxes at the corporate level only. First, the structure presented in Figure 1 is assumed without a controlling company. Then, we include a chain of ownership as well as an alternative source of finance, debt borrowed by the affiliate from external lenders.

Table 1 summarises the most important model assumptions of our calculations. The model covers the most relevant tax provisions of the national tax systems. With respect to corporate taxation, it considers headline statutory corporate profit tax rates as well as surcharges and some other special rates for particular types of income and expenditures. It also takes into account the most important features of non-income taxes, and it generally assumes a level of corporate profits and capital at which the top-bracket statutory tax rates apply. With regard to the definition of the taxable base, it considers the relevant rules with respect to depreciation and amortisation allowances, valuation of inventories and interest deductibility in case of debt financing.

**Table 1**  
Summary of the most important assumptions

Assumption with regard to ...	Value
Legal Form	Corporation
Industry	Manufacturing sector
Assets (weight)	Industrial buildings, intangibles, machinery, financial assets, inventories (at equal weights)
Sources of finance (weight)	Retained earnings (1/3), new equity (1/3), debt (1/3)
True economic depreciation	Declining balance method Industrial buildings 3.1% Intangibles 15.35% Machinery 17.5%
Real interest rate	5%
Pre-tax real rate of return (for calculation of EATR)	20%
Inflation rate	2%

**International comparison of the effective tax burden of companies**

Basically, international differences in company taxation can influence decisions on the location of real investment (investment-shifting) and decisions on the location of profit declaration for tax purposes (profit-shifting). First, we focus on tax effects on in-

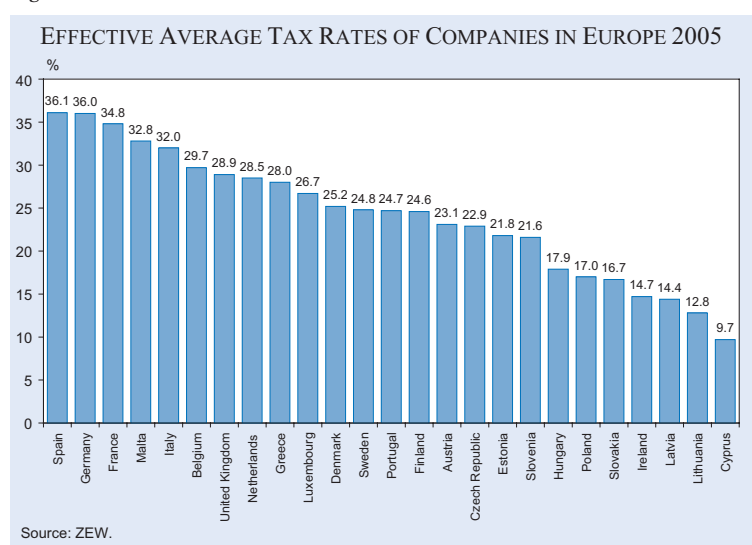
vestment decisions. For this purpose, the non-tax assumptions of the model are fixed at an equal level. The most relevant case of investment decisions of multinational groups are decisions on profitable investment projects. As described above, the EATR is the relevant measure to calculate the tax burden of a profitable project. The significance of the EATR as a relevant tax indicator has been tested empirically by Devereux and Griffith (1998), and Büttner and Ruf (2004). Therefore, we focus on an international comparison of the EATR as an indicator of the attractiveness of countries in case of location decisions. Figure 2 contains an international ranking of the EATRs companies located in the EU member states have to face. The set of results was calculated with the model presented above. A chain of ownership was not assumed. The calculations considered the tax law as of 2005. However, the effective tax burden varies significantly between each municipality, in particular in France and Germany. In Spain, a remarkable variation is also caused by a local business tax which depends on local and additionally on sectoral factors. Therefore, results were calculated by assuming an average tax level if tax levels vary due to local taxes.<sup>3</sup>

Obviously, there are substantial differences in terms of the EATR between European countries. In the EU, the highest EATRs on investments can be observed in Spain, Germany and France. In contrast, a lot of countries, especially in Eastern Europe, offer remarkably low levels of company tax rates. Moreover, small countries near the border of the European Union like Ireland, Cyprus, and the Baltic states have remarkably attractive company taxation. Nevertheless, a lot of Western European countries display a moderate effective tax burden. Countries like Spain, Germany, France and Italy, which levy significant additional local taxes, display high effective tax burdens. For example, the EATR of German companies would be at the comparatively moderate level of 23.8 percent if the local trade tax were not considered.

As a result of the comparison of the EATRs and the statutory tax rates presented in Table 2 it is obvious, that the ranking in terms of the EATR is mainly caused by the ranking of the statutory profit tax rates. This is due to the fact that the higher the expected profit rate the more the statutory tax rate influences the effective tax rate. This relationship be-

<sup>3</sup> See IBC Taxation Index 2005 (www.bakbasel.com) for regional results.

Figure 2



tween effective tax rates and the statutory profit tax rate can also be explained intuitively. If we consider a profitable investment with the same level of expenses as a marginal investment, but now accompanied by a higher level of income, the additional income is regularly taxed at the statutory tax rate without triggering additional allowances.

For a detailed discussion of tax drivers, additional results of a marginal investment expressed by EMTRs are presented in Table 2. The impact of the statutory profit tax rate becomes obvious by comparing different levels of profitability. Looking at Austria for example, the EMTR indicates, that the definition of the taxable income and non-income taxation are more favourable in other regions and do not primarily account for the low EATR in Austria. The comparatively moderate statutory profit tax rate constitutes the main reason for the good ranking of Austria compared to other Western European countries with similar or more favourable tax bases but higher statutory profit tax rates (e.g. Luxembourg, Belgium). Those European countries having the highest combined statutory tax rates on profits (Spain and Germany) also display the highest effective tax burden. Despite a lower statutory profit tax rate, the EATR for France is comparably high. This is due to the relatively high level of non-income taxes on industrial buildings and machinery, which also accounts for the highest EMTR.

Regarding the Eastern European countries, one can see that they exhibit low statutory profit tax rates and favourable rules concerning depreciation allowances. The Hungarian EATR is, in spite of a lower

combined statutory profit tax rate, higher than effective tax rates in other countries. This is due to the fact that the rules determining the tax base are less favourable in Hungary, and depreciation allowances and interest expenses do not reduce the base of the Hungarian local business tax. The comparatively high Hungarian EMTR indicates that the impact of the local business tax on the effective tax burden increases if the assumed profitability of the investment declines. The EATRs of Latvia and Lithuania only differ slightly from each other, since these countries have the same

combined statutory profit tax rates. Most notably, a variation can be found in the level of real estate taxes and the definition of taxable income, which is also reflected in a stronger divergence of the EMTRs. In Slovenia, the tax system provides favourable asset depreciation allowances so that the higher statutory profit tax rate does not become relevant. Even without the absence of any non-income taxes, these tax provisions result in a relatively low EMTR.

Table 2  
Tax rates in percent, 2005

	Statutory profit tax rate	EMTR	EATR
Austria	25.0	18.9	23.1
Belgium	34.0	20.4	29.7
Cyprus	10.0	8.8	9.7
Czech Republic	26.0	15.6	22.9
Denmark	28.0	19.1	25.2
Estonia	24.0	16.8	21.8
Finland	26.0	21.7	24.6
France	34.9	34.7	34.8
Germany	39.4	30.0	36.0
Greece	32.0	19.0	28.0
Hungary	17.7	18.6	17.9
Ireland	12.5	14.4	14.7
Italy	37.3	22.7	32.0
Latvia	15.0	12.7	14.4
Lithuania	15.0	6.9	12.8
Luxembourg	30.4	18.3	26.7
Malta	35.0	28.8	32.8
Poland	19.0	11.9	17.0
Portugal	27.5	18.4	24.7
Slovakia	19.0	10.7	16.7
Slovenia	25.0	13.2	21.6
Spain	39.9	29.2	36.1
Sweden	28.0	19.1	24.8
Netherlands	31.5	22.2	28.5
United Kingdom	30.0	26.7	28.9

Source: ZEW.

With respect to the source of finance, we found a general pattern in the countries considered: Since interest payments are completely or at least partly deductible at the corporate level, debt is tax privileged compared to the other two sources of finance. Furthermore, the effective tax burden at the corporate level on investments financed with retained earnings does not differ from those on investments financed with new equity, since assessed tax systems treat retained earnings and new equity equally at the corporate level. The only exemption is Estonia where retained earnings are tax exempt. Taxation of the marginal shareholder of a company determines the value of the firm and should be taken into account in management decisions. Nevertheless, ignoring taxes at the shareholder level is an adequate method if managers do not know the tax position of their marginal shareholder. However, domestic shareholder taxation does not affect corporate investment decisions of multinationals when there is substantial international capital mobility. Therefore, shareholder taxation is not considered in this study.<sup>4</sup>

Since it is important to calculate the effective tax burden from the point of view of the relevant decision-level, the model can be extended by adding a controlling company. This extended structure is shown in Figure 1. The consideration of the tax burden at the level of the controlling company displays the relevant effective tax burden while choosing a foreign investment location. As an example, we present the effective tax burden in terms of the EATR relevant for an investment project of German controlled affiliates within Europe. We assumed equal weighted sources of finance (retained earnings, new equity, debt) which are granted by the German holding company.

A ranking of EATRs of German-controlled affiliates is presented in the first column of Table 3 in the last part of this study. These EATRs were calculated at the level of the German controlling company. The ranking equals the ranking of the EATRs calculated at the affiliate-level, since Germany applies a system of limited exemption to domestic and foreign inter-company dividends.

<sup>4</sup> For a discussion of these issues see European Commission (2002, 142–43) or Lammersen and Schwager (2005, 69–70).

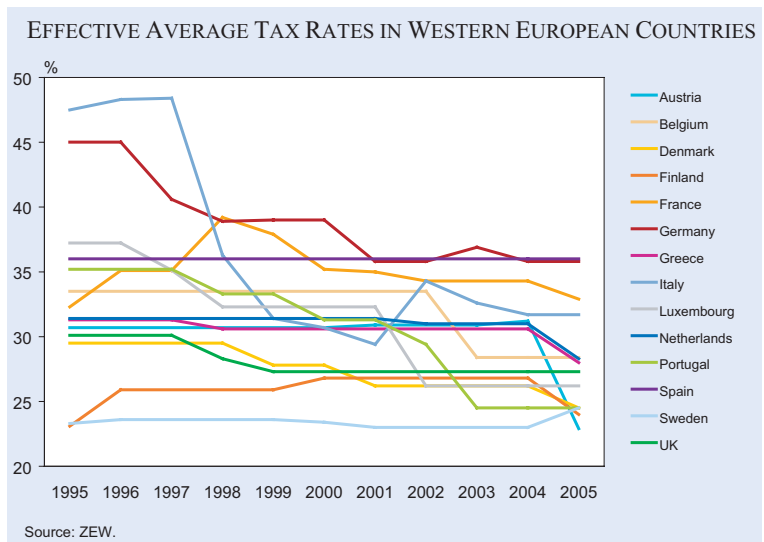
The levels of tax burden are slightly higher due to the additional German tax of five percent on inter-company dividends received and due to the higher German taxes burdened on inter-company interests. Therefore, the results calculated at the affiliate-level provide a good ranking if the controlling company's country exempts foreign dividends.

It is apparent that the differences in terms of the EATR are determined mainly by different statutory profit tax rates and that they are less determined by differences in determining tax bases. The decrease of the statutory tax rate can be named as a favourable tool to reduce the EATR and to attract highly profitable real investments. The impact of the statutory tax rate is also evident on the ranking in terms of the EMTR. Consequently, convergence of statutory tax rates would reduce differences in effective tax burden within Europe remarkably.

**Time series of effective average tax rates**

The international mobility of capital has led to competition between countries intent on attracting real investments and tax bases. Figure 2 presents the current picture of the tax competition within Europe. We have calculated time series of the EATR of EU countries to investigate changes in the European tax competition over time. A permanent trend of declining effective tax rates in Western European countries is displayed by Figure 3 for the period from 1995 to 2005. Even longer time series from 1984 on have been provided by Schreiber and Overesch (2005). During the late 1980s and the earlier 1990s, there were significant reductions in the effective tax

**Figure 3**



rates in Europe. In particular, the Scandinavian countries significantly lowered their effective tax burden by introducing dual income tax systems, which levied a lower tax burden on capital profits.

Apparently, since 1995, rate cutting activities have continued and the Western European countries have, in general, lowered effective tax rates. The average of the EATRs in the former 15 EU member states has declined from 32.1 percent in 1995 down to 27.8 percent in 2005. The highest decreases were observed in Germany and Italy. Most of the European countries have significantly decreased their statutory tax rate. Additionally, non-income taxes at the corporate level have been abolished. Furthermore, the new EU member states have joined the tax competition process. Figure 4 presents EATRs of companies located in the new EU member states from 1995 on. The effective tax rates of the new EU member states before joining the EU must be interpreted carefully, however, because they do not reflect the remarkable tax incentives like tax holidays granted by these states before EU enlargement. Therefore, the interpretation of the observed decreases in terms of the EATR of the new member states must take into account the broadening of their tax base. This was done by the abolishment of various tax holidays and investment tax credits.

During the last two years, a lot of European countries significantly lowered their effective tax rates, in particular by reducing their statutory tax rates. In 2004, Poland and Slovakia e.g., lowered their statutory tax rates from 27 percent and 25 percent, respectively, down to 19 percent. The process of tax rate reductions was continued by the Czech Republic and

Estonia in 2005. Moreover, remarkable rate cuts were observed in the former 15 EU member states. In 2005 Austria, Denmark, France, Finland, Greece and The Netherlands cut their statutory company tax rates. Thus, further reductions of tax rates can be expected.

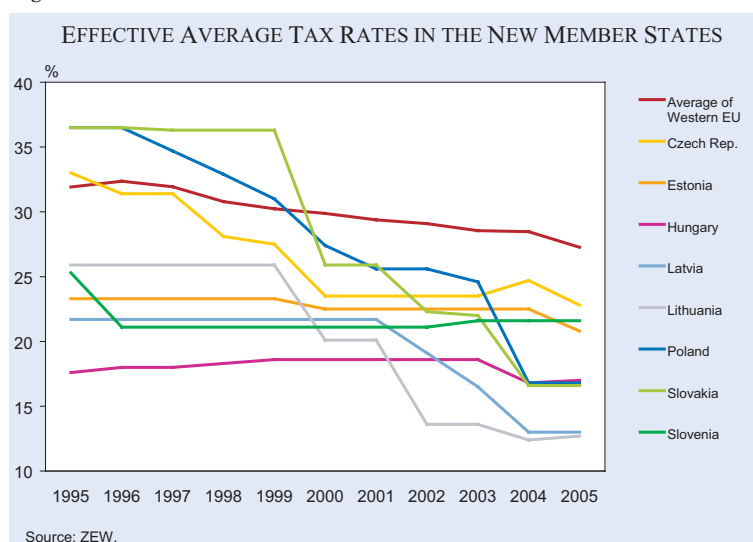
### Tax planning strategies

The presented indicators of the tax burden on investment projects point out the potential impact of taxation on decisions with respect to the location of real investments. However, in addition to the incentive to shift real investments towards a tax attractive location, a multinational company can react to international tax differences by cross-border tax planning strategies, i.e. shifting the declaration of profits for tax purposes. For example, a multinational corporation can shift profits between affiliates via transfer pricing or financial strategies. The incentive to shift profit declaration is a direct result of differences between the statutory tax rates. The first column in Table 2 shows the considerable variation of statutory tax rates within the EU.

It is even rather difficult to quantify precisely the effects generated by profit-shifting. However, the general effect of profit-shifting strategies on the effective tax burden can be analysed by the model applied in this study. Therefore, the model assumption on equally weighted sources of finance is given up in order to show effects of tax planning via finance. Basically, debt finance is tax efficient because of the tax shield generated by the interest deduction in tax accounting. Tax savings caused by interest deduction

are a function of the avoided tax rate. We present results of a German parent corporation. Since interest deductions are limited to 50 percent for purposes of the trade tax, the tax rate avoided is 32.9 percent in Germany. Consequently, in view of the comparatively high German statutory tax rate, the tax efficient strategy of a German parent company is to borrow debt. Hence, we calculated tax burdens in terms of the EATR using a debt-financed German parent company with an equity-financed affiliate in Europe that carries out the real invest-

Figure 4





ment. In Table 3, the effects of such tax-efficient financial strategies can be seen. The differences between these EATRs and those which were calculated for the mixed financing of a German controlled affiliate show the significantly positive effect of this kind of tax arbitrage. The multinational group generates high tax savings in the high tax country Germany, since it has to pay comparatively low taxes in most of the analysed countries. This behaviour results in a higher tax base for the involved low tax country, while the involved high tax country is losing tax base because of the company's interest deduction from other taxable profits generated in that country.

**Table 3**  
**EATR in percent of a German controlled affiliate financed by ...**

	Equally weighted sources of finance both in Germany and the affiliate	External debt in Germany and by equity of the affiliate	External debt borrowed by the affiliate
Austria	26.4	15.6	18.5
Belgium	32.9	23.0	22.9
Cyprus	13.2	0.6	8.7
Czech Republic	26.2	15.5	18.0
Denmark	28.5	18.0	19.8
Estonia	25.1	8.7	20.1
Finland	27.9	17.1	19.7
France	37.9	28.1	27.6
Germany	36.0	28.4	28.4
Greece	31.2	21.1	21.6
Hungary	21.2	9.4	15.3
Ireland	18.2	6.2	12.5
Italy	35.1	25.2	25.3
Latvia	17.8	5.8	12.1
Lithuania	16.3	4.3	10.6
Luxembourg	29.9	19.7	20.7
Malta	35.9	26.2	25.7
Poland	20.4	8.8	13.8
Portugal	28.0	17.4	19.4
Slovakia	20.1	8.5	13.5
Slovenia	24.9	14.1	17.0
Spain	39.1	29.4	28.9
Sweden	28.0	17.5	19.6
Netherlands	31.7	21.6	22.3
United Kingdom	32.1	21.7	23.0

Source: ZEW.

As shown above, the effective tax burden of an investment project can be reduced remarkably by using debt. In general, a multinational group decides on assignment of debt. The last column of Table 3 presents EATRs if the project is financed by locally borrowed external debt instead of external debt borrowed by the German controlling company. Equal non-tax constraints in each country were assumed in order to focus on the tax effects of this financial

strategy. A comparison of the resulting effective tax burden reveals that it is more favourable to finance an investment via German external debt instead of local external debt. Only in case of affiliates in Belgium, France, Malta or Spain, local borrowing seems to be more favourable from a tax perspective. This result is caused by the comparatively high German statutory tax rate, which can be avoided by interest deductions in Germany. From the point of view of the involved countries, there is a strong tax incentive for multinational groups to allocate external debt in affiliates located in high tax countries like Germany. Correspondingly, companies situated in high tax countries will usually have comparatively low taxable profits and therefore low effective tax payments.

In case of a controlling company in a low tax country, the effects on the location of the tax payments of the multinational group are similar. The affiliates located in high tax countries are leveraged by external or inter-company debt. The inter-company interest payments reduce the tax base in the high tax country and are taxed at the lower level of the lending company. Anti-abuse provisions of national tax law, e.g. thin-capitalization rules, are limited, since the European Court of Justice bans national regulations which can be qualified as restrictions of the freedom of establishment or the movement of capital. Moreover, the EU council directive on interest and royalties and the parent-subsidiary directive extensively prohibit withholding taxes on interest, royalties, and dividend payments to affiliated companies.<sup>5</sup> Basically, the differences in the EATRs as displayed in Figure 2 influence location decisions on real investments to a greater extent if limitations of profit-shifting exist. Thus, in particular, small and medium-sized enterprises without any foreign affiliates cannot evade the high effective tax burden of their location. It seems that multinational groups can decrease the effective tax burden of their affiliates in high tax countries via tax planning strategies.

**Conclusion**

The measurement of the effective tax burden on companies in this study relies on an approach introduced by Devereux and Griffith (1999) which fulfils all requirements to analyse company taxation as a location factor. The calculated tax burden in terms of the EATR is the relevant measure for the interna-

<sup>5</sup> See EU Directives 2003/49/EC and 2003/123/EC.

tional comparison of tax burden on profitable investment projects which are typical for international location decisions. The comparison of the EATRs indicates remarkable differences between the EU member states in 2005. Spain, Germany and France can be identified as high tax countries within the European Union, whereas Ireland, Cyprus and the new member states in Eastern Europe can be described as low tax areas. Time series of EATRs indicate a general trend towards cutting the effective tax rates in Europe. In 2005, this trend was continued and six of the former 15 EU member states lowered the tax burden of companies.

In addition to the analysis of company taxation as a location factor, the model can show the incentive to shift profits into low tax countries. Multinational groups can reduce the effective tax burden via tax planning strategies. On the contrary, small and medium-sized enterprises without foreign affiliates cannot evade the high effective tax burden of their location. Basically, differences in the effective tax rates calculated without considering special tax planning strategies influence decisions on real investments to a greater extent if constraints of profit-shifting exist.

From the point of view of each country, cutting the statutory profit tax rate seems to be a favourable strategy, because that strategy improves the position in the international tax competition on both real investments and mobile taxable profits. A reduction of statutory tax rates in high tax countries significantly lowers differences in the EATR between European countries. The impact of differences in the determination of tax bases seems to be less relevant for location decisions. Furthermore, the incentive to shift profits into another jurisdiction is directly decreased by reducing the statutory tax rate.

## References

- Büttner, T. and M. Ruf (2004), "Tax Incentives and the Location of FDI: Evidence from a Panel of German Multinationals", *ZEW Discussion Paper* 04-76.
- Devereux, M. P. and R. Griffith (1998), "Taxes and the Location of Production: Evidence from a Panel of US Multinationals", *Journal of Public Economics* 68, 335-67.
- Devereux, M. P. and R. Griffith (1999), "The Taxation of Discrete Investment Choices", *IFS Working Paper* W98/16, revision 2.
- European Commission (2002), "Company Taxation in the Internal Market", *Commission Staff Working Paper* COM (2001) 582 final.
- European Communities (2005), *Structures of the Taxation Systems in the European Union – Data 1995-2003*, [http://europa.eu.int/comm/taxation\\_customs/taxation/gen\\_info/economic\\_analysis/tax\\_structures/index\\_en.htm](http://europa.eu.int/comm/taxation_customs/taxation/gen_info/economic_analysis/tax_structures/index_en.htm) (accessed 28 October 2005).
- King, M. A. and D. Fullerton (1984), *The Taxation of Income from Capital*, University of Chicago Press, Chicago.
- Lammensen, L. and R. Schwager (2005), *The Effective Tax Burden of Companies in European Regions*, ZEW Economic Studies, Physica, Heidelberg/New York.
- Schreiber, U., C. Spengel and L. Lammensen (2002), "Measuring the Impact of Taxation on Investment and Financing Decisions", *Schmalenbach Business Review* 54, 2-23.
- Schreiber U. and M. Overesch (2005), "Effektive Steuerbelastung der Unternehmen und Steuerpolitik", *Wirtschaftsdienst* 85, 220-25.
- Sørensen P. B. (2004), "Measuring Taxes on Capital and Labor: An Overview of Methods and Issues", in P. B. Sørensen, ed., *Measuring the Tax Burden on Capital and Labor*, CESifo Series, MIT Press Cambridge, Mass.

## EDUCATION POLICY IN THE UK

STEPHEN MACHIN\* AND  
ANNA VIGNOLES\*\*

### Introduction

Throughout the post-war period there have been many attempts to reform the UK education system, often with an explicit intention to try and make it more productive. The list of education policy reforms that have been attempted over the last 50 years is quite extensive, and recently the UK (and in particular England and Wales) has introduced many innovative market-oriented reforms to its education system, in an attempt to raise standards. The most striking recent reform is that parents have increasingly been given much more choice in terms of the school attended by their children, and schools have been forced to be more accountable. Other notable reforms include a nationally prescribed curriculum, vigorous attempts to raise participation in post-compulsory schooling and the introduction of tuition fees for higher education.

In this article we describe a number of the more important educational reforms that have been introduced in the UK during the last decade or so, and where possible we provide evidence of their impact. However, in the UK, although economists and others are increasingly able to inform policy-makers on the impact and efficacy of specific policy interventions, the evaluation of specific policies in a rigorous manner unfortunately remains relatively unusual. This is mainly because the design of policy interventions is often such that they are not amenable to economic evaluation. Here we highlight what one can view as robust findings on each policy intervention, and point to where more evidence is needed.

\* Stephen Machin is Professor of Economics at University College London and Director of the Centre for the Economics of Education and Research Director at the Centre for Economic Performance, London School of Economics.

\*\* Anna Vignoles is senior lecturer at the Institute of Education, University of London and a Deputy Director at the Centre for the Economics of Education.

### Market reforms

#### *The problem*

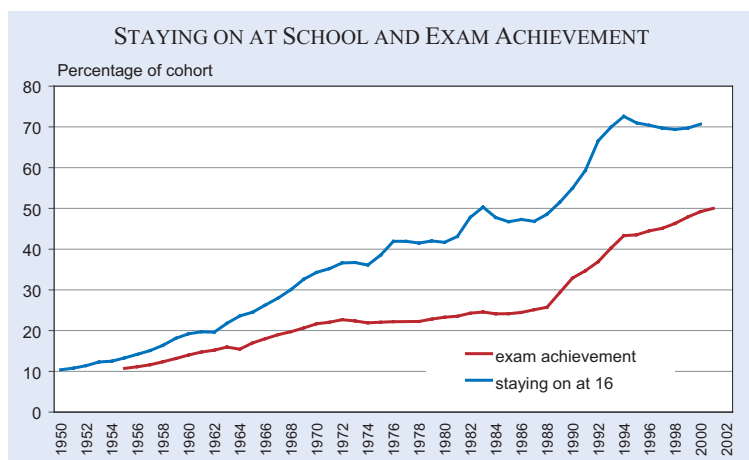
In addition to concerns about widening access and educational inequality, in the 1980s there emerged widespread fears about poor and falling standards in UK education. Specifically there were concerns that too many individuals were leaving school too early and with little in the way of basic skills.

The data is supportive of these concerns. Firstly, although the staying on rate at age 16 (the compulsory school leaving age in the UK) had been increasing over a number of years, as shown in Figure 1, it remains low by international standards.<sup>1</sup> Secondly there was growing concern that achievement had stagnated in schools, particularly in the compulsory phase. This is illustrated in Figure 1 below which shows the exam success rate at age 16, i.e. the proportion of the cohort achieving the equivalent of 5 or more grades A\*-C at GCSE.<sup>2</sup> The proportion succeeding in their examinations at age 16 remained stagnant from around 1970 to the mid 1980s. Thus in the 1980s not only were around half the cohort leaving full time education altogether after the age of 16 but they were leaving with no qualifications. More than two thirds of the cohort did not achieve examination success at age 16 and therefore entered the la-

<sup>1</sup> See, *inter alia*, OECD (2005).

<sup>2</sup> The General Certificate of Secondary Education (GCSEs) are examinations taken at the age of 16. They replaced Ordinary (O) levels and Certificates of Secondary Education (CSEs) in 1988. Here the two sets of qualifications are equivalized, thus a GCSE grade A\*-C is equivalent to an O level grade A-C or CSE grade 1.

Figure 1



Notes: Staying on is defined as the percentage of pupils staying on after the compulsory school leaving age. The exam achievement series measures the percentage of school-leavers achieving five or more higher grade GCSE (or O level) passes. Data for 1994–2000 comes from DfES Statistical Bulletins. Be-fore 1994, data are taken back using a series very kindly provided by Duncan McVicar (see McVicar and Rice 2001 for details).

Source: Clark, Conlon and Galindo-Rueda (2005).

bour market with no academic qualifications at all. Of course many of these individuals went on to take vocational qualifications, which are discussed later in this article, but nonetheless there was a widespread perception amongst UK education policy-makers that the UK had a particular problem with its so-called “long tail of low achievement”.

### *The policy*

In the light of these concerns, successive Conservative governments in the 1980s and 1990s increased the pace of reform and introduced so called “market mechanisms” into the UK education system, in an attempt to force schools to raise standards. The move towards a “quasi-market” in education was kick started by a significant piece of legislation – the 1988 Education Reform Act – which not only introduced the market reforms discussed here, but also the National Curriculum described in Section 3.<sup>3</sup>

The package of market-oriented reforms aimed to increase parental choice and thereby improve the accountability of state funded schools. Parents could, at least theoretically, choose which school their child attended and could also have representation on school governing bodies. School funding became more closely linked to student enrolment numbers, giving schools the incentive to attract and admit more students. Some schools were also allowed to take control of their own budgets<sup>4</sup> and be financed directly from the central government (as opposed to being under local government control). This gave them greater autonomy in their operations and in particular over which students they admitted to the school.

Alongside greater parental choice, policy-makers also endeavoured to improve the information available to parents about the effectiveness of schools, by way of publicly available test score information. This information was quickly re-produced by the media in the form of educational “league-tables”, showing the position of schools relative to one another, in terms of their examination success rates at age 16. Later, as more test score information became available these league tables became more sophisticated, focusing on

a range of outcome measures, rather than just examination success at age 16 and often taking a value added approach, i.e. taking account of the prior achievement of children entering a particular school. Nonetheless even today, newspapers still focus most on the overall exam pass rate in different schools as being the issue of primary interest to parents.

There are, however, significant limits to the operation of a quasi-market in the UK education system. Schools are generally not allowed to go “bankrupt”, i.e. exit from the market, and many parents still lack full information on the quality of schools. This of course weakens the incentive for schools to improve. In fact understanding the exact nature of the incentives faced by schools is a problematic area, from a theoretical perspective. The literature on public sector service delivery (Dixit 2002; Besley and Ghatak 2003) would suggest that it is not clear what the objectives of decision makers in schools actually are. Schools are not like private sector firms where the objective is generally to maximise profits. Rather, in the case of schools, teachers and head-teachers have often-conflicting objectives. Of course there are also multiple outputs from the education system, ranging from improving test scores to engendering a love of learning. Thus, as Besley and Ghatak (2003) state, the critical issue facing policy makers is to work out the best means by which competition, incentives and accountability can be brought together to enhance educational outcomes in the broadest sense. Whether the UK achieved this is, of course, an empirical question.

### *The evidence*

Bearing this theoretical literature in mind, it is unsurprising that a major concern in the UK is the unforeseen incentive effects of the market reforms. The evidence (mainly from the United States, e.g. Hoxby, 2000, 2003a, 2003b) shows that increased competition among schools and moves to decentralize school finance can enhance attainment, but can raise inequality because richer parents are better able to take advantage of a more market-oriented system. This, of course, has a productivity cost associated with it, in that often more able pupils from poor economic and social backgrounds fall behind. This is particularly important in the UK context with its tail of poor achievers, which is most obvious amongst poor and disadvantaged students. Empirical evidence is emerging that these concerns are manifest on the ground. For example, high socio-economic groups appear to

<sup>3</sup> See Le Grand (1991, 1993) or Adnett and Davies (2002) for descriptions of the “market-led” reforms in the UK.

<sup>4</sup> The major provisions of the Act were to set up a National Curriculum, to introduce testing and league tables, to offer local management of schools and to increase accountability (through a regular inspection regime and from changing the nature of school governing bodies). The Act also set up grant maintained (GM) schools that were allowed to select up to 10 percent of their pupils on the basis of ability or aptitude, and City Technology Colleges (CTCs), the first attempt to bring the private sector closely in to the state sector as they are part funded by private sector business.

have better information on, and understanding of school performance, via league tables (West and Pennell 1999). If wealthier parents act on this information, choosing for their children to attend the best schools, then there is a clear tension between strategies to raise standards and policies to reduce inequality. Socio-economic background also relates to school quality and pupil performance via peer groups. For example, attending a school with very few children from lower socio-economic groups is highly beneficial academically speaking (Feinstein 2003). If parental choice leads to greater socio-economic segregation across schools, such peer group effects will further reinforce socio-economic disadvantage.

The evidence on the extent of educational inequality in the UK somewhat counters this rather pessimistic view, however. Table 1, for example, shows staying on rates at age 16, broken down by parental income group for a number of different cohorts (the cohort birth year and the year that the cohort were aged 16 are given in the first column). Since the staying on decision occurs at age 16, the three years we look at are 1974 (for the 1958 cohort), 1986 (for the 1970 cohort) and 1996 (for the 1980 cohort). The Table shows the proportion staying on beyond age 16 for people from high, middle and low-income families. The last two columns then give the unconditional and conditional gaps in the staying on rate between children from the highest and lowest income families. The unconditional gap is simply the gap in participation rates between high and low income children. The conditional gap is the gap that re-

mains once one has allowed for some other factors that influence participation, in particular ability of the child and gender.

Clearly there has been a rise in the staying on rate for all children, regardless of the income level of their family. Furthermore, in the early period inequality grew. Between 1974 and 1986, staying on rates rose fastest for children from high-income backgrounds. By contrast, during the period when some of the market reforms discussed above were in place, i.e. between 1986 and 1996, staying on rates grew faster for children from the lowest income backgrounds. Thus educational inequality rises between the first two cohorts, by 0.14 percentage points, and falls, by 0.13, between the second and third. This provides very preliminary evidence on the impact of market reforms on inequality however, given that the sixteen year olds considered in 1996 had spent most of their time in the education system before market reforms were introduced.

Of course the motivation behind the introduction of the market reforms was to raise standards and achievement, rather than issues related to inequality. The evidence on the impact of the reforms on children's achievement is minimal however. Empirical evidence from the US (Chubb and Moe 1990) is supportive of the view that decentralized schooling systems can produce better results, measured in terms of educational outcomes (see also Hoxby 2000). The only evidence for the UK to date is Bradley et al. (2001) which found that schools with the best examination

performance grew most quickly and that increased competition between schools led to improved exam performance. More recent work finds only very limited evidence of choice and competition effects on pupil achievement. Gibbons, Machin and Silva (2005) report little evidence of a link between choice and achievement, but find a small positive association between competition and school performance. However, they attribute this to endogenous school location or pupil sorting. Only in a minority of cases, the one in five or so of the school population who attend religious primary schools is there any positive causal impact of competition on pupil achievement.

**Table 1**  
**Staying on rates (proportions) by parental income group**

Cohort and year	Parental income group			Educa-tional inequality (uncondi-tional)	Educa-tional inequality (condi-tional)
	Lowest 20 percent	Middle 60 percent	Highest 20 percent		
1958 cohort (1974)	.21	.27	.45	.24 (.02)	.24 (.02)
1970 cohort (1986)	.32	.43	.70	.38 (.02)	.39 (.02)
1980 cohort (1996)	.61	.71	.86	.25 (.03)	.23 (.03)
Change 1974–1986	.11	.16	.25	.14 (.03)	.15 (.02)
Change 1986–1996	.29	.28	.16	-.13 (.04)	-.16 (.04)
Change 1974–1996	.40	.44	.41	.01 (.04)	-.01 (.04)

Notes: Sample sizes are 5706 for the NCDS 1958 cohort, 4706 for the BCS 1970 cohort and 1610 for the BHPS 1980 cohort. The conditional model adds controls for family size, sex, parents' age and living in a single-parent family. Educational inequality in the conditional case is a marginal effect derived from a probit model of staying on beyond 16 including dummy variables for quintiles of family income. This marginal effect is defined as  $\Pr[\text{Stay On} | \text{Top Income Quintile}] - \Pr[\text{Stay On} | \text{Bottom Income Quintile}]$ . Standard errors in parentheses.

Source: Blanden, Gregg and Machin (2005), Table 5.3.



One might of course argue that Figure 1 shows clear evidence that the totality of reforms introduced during the late 1980s and early 1990s had a positive effect, particularly on achievement at age 16. Certainly there has been a dramatic rise in the examination success rate since the late 1980s. However, some care is needed before interpreting the data on this. In 1988 there was a reform of the examination system at age 16, with a switch from the GCE ‘O’ level system to the introduction of the GCSEs. This reform moved the education system from one that rationed the number of O level passes in a given year to one where, at least in principle, everyone could pass a GCSE (see Gipps and Stobart 1997; Blanden, Gregg and Machin 2005). Furthermore, the GCE ‘O’ level system was purely exam based, whereas GCSEs often have a substantial coursework component. It has been argued (see Kingdon and Stobart 1998) that this also facilitated an increase in the pass rate achieved.<sup>5</sup> Certainly the most dramatic feature of the data is the structural break that occurs in examination achievement in 1988, with substantial rises in achievement from 1988 onwards. Proving causality however is not possible and this illustrates the problem of evaluating the impact of nationally introduced education policies.

**Curriculum reform**

*The problem*

In addition to concerns about participation and examination achievement, it has also been recognised by education policy-makers that the UK has a particular problem with basic skills. Table 2 shows the level of basic skill in numeracy and literacy by age group, taken from the International Adult Learning Survey (IALS 1995). Specifically, the table shows the proportion of adults with numeracy and literacy skill levels at IALS Level 2 or above (deemed by policy-makers to be the minimum level of skills required to

**Table 2**  
**Numeracy and literacy performance by age from the international adult literacy survey**

	Percentage of adults at IALS level 2 or above					
	Numeracy			Literacy		
	Age 16–25	Age 26–35	Age 36–45	Age 16–25	Age 26–35	Age 36–45
Belgium (Flanders)	93	91	83	92	88	80
Switzerland (German)	93	87	81	93	83	76
Netherlands	92	93	90	92	94	91
Sweden	95	96	93	96	95	93
Germany	96	95	94	91	88	86
Ireland	82	80	77	84	84	79
Britain	78	80	81	83	82	83
USA	74	80	82	77	80	81

Notes: Based on IALS measurement of “quantitative literacy” and “prose literacy”.

Source: OECD, *Literacy, Economy and Society*, pp. 152 and 154.

function effectively in the labour market). Amongst middle age workers, age 36-45, the UK performs around the average, as compared to other IALS countries. However, unlike in other countries, in the UK numeracy and literacy skill levels actually start to fall among younger workers. This was suggestive that the UK education system was becoming increasingly less effective in producing workers with adequate levels of basic skills.

*The policy*

To tackle the problem of poor literacy and numeracy, as well as address general concerns about poor standards, the UK introduced two other significant national policies. Firstly, in the late 1980s a standardized national curriculum was introduced for pupils aged between 7 and 16. The purpose of the national curriculum was to raise standards by ensuring that all students study a prescribed set of subjects up to a minimum level until the age of 16. The second policy reform, in 1998, was the introduction of the National Literacy and Numeracy Strategies. These strategies involved all primary schools allocating part of the daily curriculum to literacy and numeracy hours, with the specific aim of developing pupils’ basic skills. The content of these daily literacy and numeracy lessons, and indeed how they should be taught, was tightly prescribed by central government. Students’ understanding of the curriculum also began to be tested, via the use of national tests taken at ages 7, 11, 14 and 16 (or Key Stage 1, 2, 3 and 4).

Whilst the national curriculum was an example of a highly centralised education policy, contrasting to the

<sup>5</sup> This may also have affected the distribution of educational attainment at age 16 (for example, the shift to coursework seems to have been a factor in the strong improvement shown by girls relative to boys: see Gorard, Rees and Salisbury (2001) or Machin and McNally (2005).

devolution of power and accountability to schools inherent in the market reforms described above, it has also generated more information for parents on the quality of each school. Thus the national curriculum and accompanying testing regime may have also enhanced the operation of the quasi-market.

To understand why education policy-makers felt the need for a tightly prescribed national curriculum and daily lesson plan in primary schools, it is helpful to consider briefly the labour market for teachers. There are currently severe problems in attracting high ability, highly qualified students into teaching in the UK (Chevalier and Dolton 2005). In the short run, it appeared that being more prescriptive in what teachers should be teaching (and teaching them how to teach it) might raise standards, at least in the absence of being able to recruit more effective teachers. Of course in the longer term, an important policy aim is to try and re-establish teaching as an important and well-respected profession, which sits uneasily with policies that take away teachers' autonomy. However, this longer run objective clearly requires policy-makers to think seriously about improving the total compensation package for teachers, including of course their pay relative to other well respected professions, as well as their non-pecuniary conditions of work (Chevalier and Dolton 2005). Whilst some reforms on teacher pay and conditions have been introduced (performance related pay for example), and others are on the legislative agenda, there is no empirical evidence on the impact of these changes since the policies have been nationally introduced with no attempt at prior evaluation. Indeed that is the case with much of the curriculum reform described here, as is evident from the discussion on evidence below.

*The evidence*

Given that the national curriculum was introduced at the national level, a robust evaluation of its impact has not been possible. There is, however, evidence on the effectiveness of the literacy hour, since it was piloted prior to national implementation. Machin and McNally (2004) undertook an economic evaluation of the National Literacy Project (NLP), which was a pilot project, where the literacy hour was introduced

in about 400 English primary schools in 1997 and 1998. This pilot resulted in children being exposed to the literacy hour in these schools for two years before the national roll out took place. Their analysis shows substantial improvements in reading and English took place (for example, reading scores rose by around .09 of a standard deviation) for a policy that is not very costly (just over £26 per pupil/year, in 2004 prices). This work suggests that improving the way in which teaching is delivered – in their case through the well-structured literacy hour – can provide a cost effective means of raising pupil attainment.

**Raising participation in post-compulsory schooling**

*The problem*

The relatively small proportion of young people staying on beyond the compulsory schooling age in the UK has already been discussed in Section 1. However, if school leavers went on to undertake part-time high quality vocational training that resulted in well-respected qualifications with high value in the labour market, then the fact that too few young people stay on in full time education would not be such a major issue. In the UK however, this is not the case. As Table 3 shows, the UK still has lower proportions of its workers at Level 2 or above, whether one considers vocational or academic qualifications. In fact the gap between the UK and other countries is higher when one focuses exclusively on vocational qualifications, particularly at Level 2. For policy-makers at least, this international evidence suggests that the UK has too few young people pursuing a vocational qualification, and too many dropping out of education and training altogether, and thereby entering the labour market with no qualifications.

**Table 3**  
**Percentages at qualification levels 2+ and 3+ in the UK, France and Germany, by type of qualification**

	Level 2+			Level 3+		
	UK	France	Germany	UK	France	Germany
16–64 year olds, general	27	31	25	20	25	22
16–64 year olds, vocational	27	41	58	17	12	52
25–28 year olds, general	33	40	33	24	36	30
25–28 year olds, vocational	28	43	52	17	18	48

Notes: The data refer to 1998, except for Germany, which is for 1997. The German results refer to the old West Germany only.

Source: Steedman (1999).

### *The policy*

There are two major policies that are worthy of mention in relation to attempts to raise participation in post-compulsory schooling in the UK. The first is the perennial (and often ineffectual) attempts at qualification reform, designed to enhance the attractiveness and labour market value of vocational qualifications. The second policy was the introduction of an Education Maintenance Allowance, which paid individuals from disadvantaged backgrounds a small means-tested allowance if they stayed on in full time education beyond the age of 16. We start however, by discussing the reform of the vocational qualification system.

Vocational education in the UK is seen as a particularly problematic area. The system of vocational training and qualifications in the UK is complex and has changed substantially over time. Certainly there is no unified system of vocational education, as is found in some other countries such as Germany. There are hundreds of different vocational qualifications currently available. Different providers offer very different qualifications, with quite different requirements in terms of achievement. This has left students, parents and employers somewhat confused about the content and economic value of different vocational qualifications.

Despite this, full-time vocational education is chosen by around 25 percent of all 16 and 17 year olds in the UK. This has risen from just under 15 percent in the mid-1980s (West and Steedman 2003). Vocational education therefore represents a sizable part of the UK education system. Part of the problem of vocational education in the UK stems from continued unsuccessful attempts to achieve so called “parity of esteem” between vocational and academic education. Numerous reforms have been introduced, all in an effort to improve the status of vocational qualifications, as compared to their academic counterparts (GCSE/A-levels). Of course the instability that this continual reform generates, itself undermines the value of vocational qualifications.

In the UK in the 1960s vocational education typically consisted of one day a week of study at a further education college, in conjunction with an apprenticeship. This system led to qualifications being provided by different awarding bodies depending on the industry. During the 1970s and 1980s the UK apprenticeship system virtually collapsed in the traditional apprenticeship sectors. Various initiatives attempted to

replace the traditional apprenticeships (West and Steedman, 2003) with yet more qualifications, increasingly taken full time at further education colleges and with no work based element. These initiatives are too numerous to mention but the most recent reforms of note are the development of National Vocational Qualifications or NVQs and the General National Vocational Qualifications or GNVQs.

Introduced in 1988, NVQs were originally intended as competence based qualifications. They were designed to certify existing occupational knowledge and skills, and were targeted at those in work. Many criticisms have been aimed at NVQs, in particular that they are too low level and do not require sufficient vocational knowledge and skill. GNVQs, on the other hand, were introduced in 1992 and were designed to be largely classroom-based taught vocational qualifications. The aim was to provide an option that would enable students to either enter the world of work directly afterwards or to continue with further study. The reforms continue apace however. Most recently vocational GCSEs have been introduced, by design equivalent to their academic GCSE counterparts, and GNVQs are shortly to be abolished.

Alongside this, there has been a concerted effort to reintroduce high quality apprenticeships. In 1995 the Modern Apprenticeship scheme was introduced to provide a high quality vocational option for more able students. Modern Apprenticeships are modelled on the German dual system of apprenticeship, and are aimed at young people (age 16-19). The apprenticeship prepares the worker for a NVQ level 3 qualification, and generally lasts around 3 years. The UK apprenticeship rate is now greater than that of France, Finland and the US, although still well below the levels in Germany and Denmark.

Policy-makers did recognise however, that simply changing the nature of vocational qualifications was not, by itself, going to necessarily encourage higher participation. In particular, it was unlikely to encourage poorer students to continue on in full-time education to take, for example, a GNVQ. The policy-question was therefore how does one target poorer students and encourage them to stay on in education for longer? The policy solution was the Education Maintenance Allowance (EMA). The objective of the EMA was to raise post-compulsory educational participation and retention in education of young people (age 16-19) from low income families. Education Maintenance Allowances are weekly term time only pay-

ments made to students aged 16-19 for staying on in full time education for 2-3 additional years. The amount paid to the student varies and is means tested, with a maximum of £30 per week. Receipt of the allowance is conditional on attendance at school or college and in addition the scheme has financial bonuses for students who progress well in their chosen education course. The EMA scheme was therefore designed to give young people from disadvantaged backgrounds an added incentive to stay on in education and to help them meet some of the additional costs associated with full time education. It is estimated that in 2004, around 50 percent of young people aged 16-19 lived in households with an income level that qualified them to be eligible for the EMA.

One unique feature of the Education Maintenance Allowance scheme is that it was thoroughly evaluated prior to full national implementation and the evaluation design was methodologically robust. In 1999, Education Maintenance Allowances were introduced in England on a pilot basis. These pilots were then extended to about a third of the country between 2000-2004. In 2004, the EMA was rolled out throughout the UK and, as we shall see when discussing evidence, it is viewed as a highly successful policy intervention.

#### *The evidence*

We start by discussing the impact of the continual reform of the UK vocational education system, i.e. the development of new vocational qualifications. Perhaps the best way to evaluate new qualifications is to consider their labour market value, since this reflects the demand by employers for this type of qualification and the skills embodied in it and will indicate the attractiveness or otherwise of the qualification for young people. The evidence on this is clear. NVQs and indeed GNVQs, have minimal economic value in the labour market (Dearden et al. 2002). In particular, NVQ2 qualifications actually have a negative impact on individuals' wages, in many sectors of work. This is suggestive of a negative signalling effect from taking these low level qualifications, i.e. employers perceive that individuals who take these vocational qualifications to be of low ability and actually prefer individuals with no qualifications at all. This is reinforced by the fact that many individuals who have been unemployed for some time take NVQ2 qualifications. Even higher level NVQ qualifications, i.e. Level 3-5, have very low labour market value and considerably lower than their academic

equivalents. Furthermore, it is not the case that all vocational qualifications have lower labour market value. Firstly, the older vocational qualifications (HNC/HND) do have substantial labour market value, arguably on a par with their academic counterparts. In terms of new qualifications, where an apprenticeship leads to a NVQ level 3 qualification, it leads to a substantial wage premium (McIntosh 2002). The evidence therefore points to the lack of success policy-makers have had, by and large, in reforming the vocational qualification system, but also indicating that some interventions, such as the re-introduction of apprenticeships, might be moving in the right direction.

Of course a major reason that employers hold vocational qualifications in lower esteem (and pay individuals with these qualifications less) is precisely because in the UK less able students choose to go down the vocational route (Clark, Conlon and Galindo-Rueda 2005). However, there are additional problems within the vocational education system itself, at least in the UK. The proliferation of vocational qualifications in the UK has led to a system little understood by employers. If employers are not even sure what a person has learned as a result of taking a particular vocational qualification, it is unsurprising that some vocational qualifications have very little or nil economic value. Continuing to develop new vocational qualifications in the fruitless struggle for parity with academic qualifications may actually exacerbate the problem.

By contrast, policy-makers have had much more success with their Education Maintenance Allowance policy. The results from the EMA evaluation suggest substantial impacts from the subsidy. Dearden et al. (2005) found that overall, educational participation post 16 was 4.5 percentage points higher amongst those eligible for the EMA (as compared to an education participation rate of 64.7 percent in the control sample). The EMA had a different impact for different groups of students however. The EMA increased initial educational participation of eligible males by 4.8 percentage points and of eligible females by 4.2 percentage points. Of course if the EMA encourages students who would otherwise drop out to stay on in education it is a possibility that these students will find their course of study too difficult. They may then drop out subsequently. However, for individuals in their second year of receipt of EMA, the impact of EMA increased to 7.6 percentage points for males and 5.3 percentage points for

females. This is encouraging in terms of the longer-term impact of EMA. Dearden et al. (2005) concluded that the EMA not only increases participation in full time education beyond the compulsory school leaving age but also enhances retention in full time education subsequently.

The evaluation by Dearden et al. (2005) also found that around half of the individuals who stayed in education were drawn from inactivity rather than from work. Thus the subsidy did not simply draw young people away from the world of work and back into education. Rather a significant proportion of the people who continued in education due to the EMA would otherwise have been inactive.

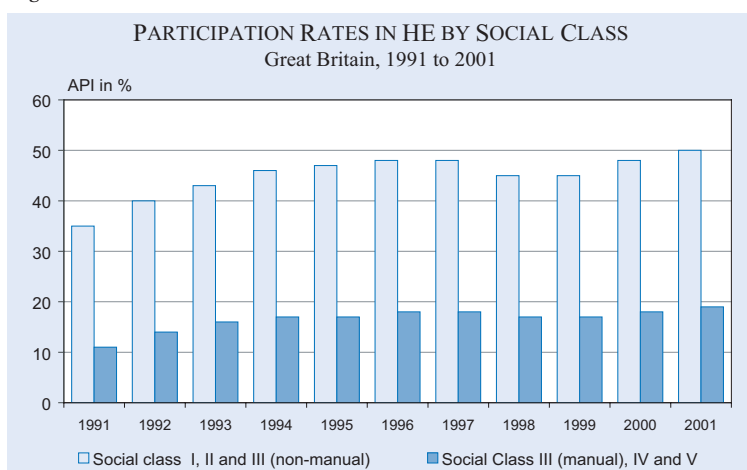
In summary, the attempts to raise educational participation post-16 have had mixed success. Reform of the vocational education system has not, by and large, been successful. This is obvious from the fact that fifteen years later, policy-makers are still undertaking radical reform of this system. However, the Education Maintenance Allowance offers a more promising way of raising participation, although it is too early to deduce its impact on participation nationally.

## Higher education reform

### *The problem*

Higher education in the UK is viewed as a success story, with continually rising participation in higher education (HE) since the late 1960s. However, there have been concerns about who is accessing HE. As shown in Figure 2, even during the last fifteen years, participation in HE has largely been the preserve of the higher socio-economic groups in the UK. Furthermore, there is evidence that the gap in HE participation between richer and poorer students actually widened, at least in the mid-and late 1990s (Machin and Vignoles 2004; Blanden and Machin 2004). Contrary to what many believed before the expansion of higher education, the expansion appears to have actually acted to increase educational inequalities, so that a greater share of HE participants is from well-off backgrounds (Machin and Vignoles 2004). This

**Figure 2**



Notes: Top 3 social classes: professional, managerial, intermediate occupations, skilled non-manual; Bottom 3 social classes IIM, IV, V – unskilled non-manual, skilled manual, unskilled manual.

Source: Department for Education and Skills Age Participation Index.

means that although poorer students are more likely to go on to higher education than they were in the past, the likelihood of them doing so relative to their richer peers is actually lower than was the case in earlier decades. This is one of the key policy challenges facing many governments and certainly is a major problem in the UK. This matters if, as appears to be the case, one is in a situation where more able children from less advantageous economic backgrounds are missing out.

### *The policy*

The government's primary HE policy over the last two decades has been to enable and encourage further expansion of the sector. Expansion is arguably needed for two reasons. Firstly, policy-makers want to continue to expand the supply of skilled labour in order to compete internationally. Secondly, they want to improve the chances of anyone, regardless of socio-economic background, attending HE. Thus the government has a target of getting 50 percent of all young people to attend university by 2010. Of course, the next issue is how one might finance such an expansion. In the UK, higher education has traditionally been free at the point of use for students. However, as higher education participation rose in the 1980s and 1990s, this became increasingly problematic. Firstly, the level of per capita resourcing in HE fell dramatically, as student numbers were expanded whilst funding remained more or less constant in real terms. For example, between 1989 and 1997 per student funding fell by 36 percent (Clark, Conlon and Galindo-Rueda 2005). In response to these problems, a means tested tuition fee was introduced in 1998. The fee was for a maximum of



£1,000 per year and had to be paid upfront, i.e. prior to the student starting the year of study in HE. Poorer students were exempt from these fees. Previously however, poorer students were also entitled to a grant to subsidise their maintenance costs whilst at university. Such grants were gradually reduced in value in real terms and phased out completely in 1999. Grants were replaced by means-tested loans, repayable on an income contingent basis after graduation.

In 2003, the Labour government proposed some further radical reforms. The purpose of the reforms was to allow universities to increase their funding, by levying higher tuition fees on students, and for institutions to differentiate themselves by charging higher or lower fees than other institutions. Universities will therefore be able to charge a higher amount, up to £3,000 per year. The fee is not payable up front however. Instead the fee debt will be paid post graduation and on an income contingent basis. In other words, graduates rather than students will pay back tuition fee loans and if their income level is sufficiently low they do not have to meet the debt payments on the loan.

#### *The evidence*

There are two main questions concerning the expansion of higher education. First, one needs to address the question as to whether more graduates are needed and whether, in the face of an increased supply of graduates, the investment in postsecondary degree acquisition remains one that yields a significant return. Secondly one needs to determine the impact of expansion, and financial reform, on inequality in HE participation.

The demand for graduates still outstrips the supply, and so there is still a significant payoff for possessing higher educational qualifications (Blundell, Dearden and Sianesi 2005; McIntosh 2005). For example, in Table 4 below, wage and employment differentials between graduates and non-graduates are shown over the period 1980-2004. It is very clear that the graduate wage premium has not fallen despite the very sharp increases in the supply of graduates and in fact has actually risen sharply at the same time as the supply rises (until the 2000s when it stays flat). Some recent evidence does suggest a slight fall in the wage premium for very recent graduates in specific subjects (Sloane and O'Leary 2004; Walker and Zhu 2005). In particular, very recent arts and humanities graduates may have seen a small fall in the immediate return to their degrees.

**Table 4**

**Aggregate trends in graduate/non-graduate employment and relative wages, 1980–2004**

	UK labour force survey/ General household survey	
	Graduate share of employment, in %	Relative weekly wage (full-time)
1980	5.0	1.48
1985	9.8	1.50
1990	10.2	1.60
1995	14.0	1.60
2000	17.2	1.64
2004	21.0	1.64
1980–2004	16.0	.16
1980–1990	5.2	.12
1990–2000	7.0	.04
2000–2004	3.8	.00

Notes: Sample is all people aged 18–64 in work and earning, except for relative wages, which are defined for full-time workers. The relative wage ratios are derived from coefficient estimates on a graduate dummy variable in semi-log earnings equations controlling for age, age squared and gender (they are the exponent of the coefficient on the graduate dummy).

Source: Machin and Vignoles (2005), updated.

Overall, however, the evidence currently supports the view that the demand for graduates is sufficient to justify further expansion, although we are probably likely to see an increasing diversity of graduate outcomes, with some lower ability graduates in certain subjects experiencing low returns to their degrees. One effect of the introduction of tuition fees, however, is that it is likely that students will be more responsive to market signals about the value of different degrees and will make their choices accordingly. This should, in the absence of other market constraints, mean that we do not see substantial expansion and over supply of certain types of graduate.

Evidence on the extent of inequality in HE participation is somewhat more complex. The gap in HE participation between richer and poorer students has actually widened in recent years (Machin and Vignoles 2004; Marcenaro-Gutierrez, Galindo-Rueda and Vignoles 2004). Further examination suggests that much of this gap is due to the fact that poorer students lack the qualifications required to enter HE. In terms of educational policy, the question is whether tuition fees have worsened the situation. Tuition fees were introduced in the UK in a manner that has prevented any robust evaluation of their impact on student participation. From an economic perspective, the extensive and robust empirical evidence of persistent high private returns to a postsecondary degree would appear to provide justification for greater student contributions in the form of higher fees. However, the critical point here seems to return to the issue of the

socio-economic mix of students who attend university. If fees are charged (which may in future be, as in the US, differential fees by subject and/or university) then it is absolutely vital that this does not act to reinforce the inequalities already present. Descriptive evidence on the impact of tuition fees introduced in 1998 indicates that the widening of the gap in participation between richer and poorer students is not as a direct impact of tuition fees, given that it occurred prior to the introduction of fees (Marcenaro-Gutierrez, Galindo-Rueda and Vignoles 2004).

### Concluding remarks

The evidence on the impact of the various reforms discussed in this article is rather patchy, although it is better for some areas of the education sequence than for others. For example, we are still a long way off from knowing what the impact of market-oriented reform is on student performance and inequality in the UK education system. But whilst we do not know what impact the National Curriculum has had in the UK on pupil achievement, we do know that the literacy hour has proved to be effective in improving primary pupil's reading skills. We know that some of the vocational qualifications introduced in recent decades (NVQ2 for example) have failed to attract any value in the labour market, suggesting these reforms have been unsuccessful in this dimension at least. We know that paying young people from disadvantaged backgrounds a relatively small allowance to stay on in school is likely to increase their chances of doing so. We know that despite expansion of HE, young people from disadvantaged backgrounds are still much less likely to go to university and more worryingly that the gap in participation between rich and poor actually widened in the 1980s and 1990s. We suggest that this is due largely to inequalities earlier in the education system, rather than financial reform and other factors in HE. In other words, poorer students are much less likely to acquire the necessary qualifications to get into HE in the first place. We have only weak evidence on the direct impact of tuition fees, but the evidence we do have suggests that they have not significantly impacted on HE participation by lower income students in the UK.

Of course knowing what works in education is not sufficient to inform policy. As economists, we need to inform policy-makers about what works and at what cost, relative to alternative policy options. Yet there remains a deficiency of good cost benefit evaluations

in the field of education. Perhaps the best example of a properly designed evaluation in the UK is the Educational Maintenance Allowance evaluation, but even this does not include a full cost benefit analysis, and there are relatively few examples in the field as a whole. The fact that developments on the cost benefit front have been markedly slow in this field is not due to the laziness of researchers however. Unlike in other fields, such as health economics, data on educational and labour market outcomes has been far more readily available than data on the myriad inputs that go into the education process, such as teachers, books, infrastructure, peer groups and parenting. This is slowly being rectified in the UK at least, with the construction of superior data sets. Therefore in the next 5-10 years, this is where one would expect to see the most progress being made in terms of empirical analysis and how it is used to inform the design and implementation of education policy.

### References

- Adnett, N. and P. Davies (2002), *Markets for Schooling*, Routledge, London.
- Besley, T. and M. Ghatak (2003), "Incentives, Choice, and Accountability in the Provision of Public Services", *Oxford Review of Economic Policy* 19, 235-49.
- Blanden, J. and S. Machin (2004), "Educational Inequality and the Expansion of UK Higher Education", *Scottish Journal of Political Economy*, Special Issue on the Economics of Education 54, 230-49.
- Blanden, J., P. Gregg and S. Machin (2005), "Educational Inequality and Intergenerational Mobility", in S. Machin and A. Vignoles, eds., *What's the Good of Education? The Economics of Education in the United Kingdom*, Princeton University Press, Princeton.
- Blundell, R., L. Dearden and B. Sianesi (2005), "Measuring the Returns to Education", in S. Machin and A. Vignoles, eds., *What's the Good of Education? The Economics of Education in the United Kingdom*, Princeton University Press, Princeton.
- Bradley, S., G. Johnes and J. Millington (2001), "School Choice, Competition and the Efficiency of Secondary Schools in England", *European Journal of Operational Research* 135, 527-44.
- Chevalier, A. and P. Dolton (2005), "The Labour Market for Teachers", in S. Machin and A. Vignoles, eds., *What's the Good of Education? The Economics of Education in the United Kingdom*, Princeton University Press, Princeton.
- Chubb, J. and T. Moe (1990), *Politics, Markets and America's Schools*, The Brookings Institution, Washington DC.
- Clark, D., G. Conlon and F. Galindo-Rueda (2005), "Post-Compulsory Education and Qualification Attainment", in S. Machin and A. Vignoles, eds., *What's the Good of Education? The Economics of Education in the United Kingdom*, Princeton University Press, Princeton.
- Dearden, L., S. McIntosh, M. Myck and A. Vignoles (2002), "The Returns to Academic and Vocational Qualifications in Britain", *Bulletin of Economic Research* 54, 249-74.
- Dearden, L., C. Emmerson, C. Frayne and C. Meghir (2005), "Education Subsidies and School Drop-Out Rates", *Centre for the Economics of Education Discussion Paper*, in press.
- Dixit, A. (2002), "Incentives and Organizations in the Public Sector", *Journal of Human Resources* 37, 696-727.
- Feinstein, L. (2003), "Inequality in the Early Cognitive Development of British Children in the 1970 Cohort", *Economica* 70, 73-97.

Gibbons, S., S. Machin and O. Silva (2005), "Choice, Competition and Pupil Achievement", *Centre for Economics of Education Discussion Paper*, in press.

Gipps, C. and G. Stobart (1997), *Assessment: A Teacher's Guide to the Issues*, Hodder and Stoughton, London.

Gorard, S., G. Rees and J. Salisbury (2001), "Investigating the Patterns of Differential Attainment of Boys and Girls at School", *British Educational Research Journal* 27, 125–39.

Hansen, K. and A. Vignoles (2005), "The United Kingdom Education System in an International Context", in S. Machin and A. Vignoles, eds., *What's the Good of Education? The Economics of Education in the United Kingdom*, Princeton University Press, Princeton.

Hoxby, C. (2000), "Does Competition Among Public Schools Benefit Students and Taxpayers?", *American Economic Review* 90, 1209–38.

Hoxby, C. (2003a), *The Economics of School Choice*, Chicago University Press, Chicago.

Hoxby, C. (2003b), "School Choice and School Competition: Evidence from the United States", *Swedish Economic Policy Review* 10, 9–66.

Kingdon, M. and G. Stobart (1998), *GCSE Examined*, Falmer Press.

Le Grand, J. (1991), *Equity and Choice*, Harper Collins, London.

Le Grand, J. (1993), *Quasi-markets and Social Policy*, Macmillan, London.

Machin, S. and S. McNally (2004), "The Literacy Hour", *Centre for the Economics of Education Discussion Paper* 43.

Machin, S. and S. McNally (2005), "Gender and Student Achievement in English Schools", *Oxford Review of Economic Policy*, in press.

Machin, S. and A. Vignoles (2004), "Educational Inequality: The Widening Socio-Economic Gap", *Fiscal Studies* 25, 107–28.

Machin, S. and A. Vignoles (2005), eds., *What's the Good of Education? The Economics of Education in the United Kingdom*, Princeton University Press, Princeton.

Marcenaro-Gutierrez, O., F. Galindo-Rueda and A. Vignoles (2004), "The Widening Socio-economic Gap in UK Higher Education", *National Institute Economic Review* 190, 70–82.

McIntosh, S. (2002), "Further Analysis of the Returns to Academic and Vocational Qualifications", *Department for Education and Skills Research Report*, no. 370.

McIntosh, S. (2005), "Evidence on the Balance of Supply and Demand for Qualified Workers", in S. Machin and A. Vignoles, eds., *What's the Good of Education? The Economics of Education in the United Kingdom*, Princeton University Press, Princeton.

McVicar, D. and P. Rice (2001), "Participation in Full-Time Further Education in England and Wales: An Analysis of Post-War Trends", *Oxford Economic Papers* 53, 47–56.

OECD (2005), *Education at a Glance*, Paris.

Sloane P. J. and N. C. O'Leary (2004), "The Return to a University Education in Great Britain", *IZA Discussion Paper* 1199, Bonn.

Steedman H. (1999), "Report to the Department for Education and Employment: Updating of Skills Audit Data 1998", *Centre for Economic Performance Working Paper*, London School of Economics.

Walker, I. and Y. Zhu (2005), "The College Wage Premium: Over Education and the Expansion of Higher Education in the UK", University of Warwick, mimeo.

West, A. and H. Pennell (1999), "School Admissions: Increasing Equity, Accountability and Transparency", *British Journal of Education Studies* 46, 188–200.

West, J. and H. Steedman (2003), "Finding Our Way: Vocational Education in England", *Centre for Economic Performance Occasional Paper*.

## WOMEN'S ATTITUDES TOWARDS FAMILY AND GENDER ROLES

One important set of theories about changes in childbearing behaviour stresses the importance of fundamental changes in women's values and attitudes towards childbearing and gender roles. Higher education attainment and labour market participation among women have fuelled the diffusion of new values – such as autonomy and financial independence – among younger cohorts of women, and greater awareness of the “incompatibility” between professional and family roles that still characterise many OECD countries. One can distinguish between “structural-role” incompatibility, i.e. between the actual opportunities available to women and the constraints that they face when trying to take advantage of these opportunities, and “cultural” incompatibility, which relates to broad ideologies, values and norms concerning the role of women in the society. Evidence is presented here on the latter and on their role in childbearing behaviour. Data from the 2000 wave of the World Values Survey are used to describe differences in attitudes to family and gender roles between two cohorts of women (those aged between 15 and 34, and between 35 and 50 in 2000).

The Table presents information on the share of women in the two age groups that agree or strongly agree with a range of statements that reflect the “traditional” role of women in families and society. Survey questions relate to whether respondents (agree that):

- i) “when jobs are scarce men should have more right to work than women”;
- ii) “marriage is not an outdated institution”;
- iii) “women need to have children to be satisfied”;
- iv) “disapprove women as lone parents”;
- v) “working mothers cannot have the same warm and stable relation with children”;
- vi) “being a housewife is as fulfilling as working in a paid job”.

For each of these six questions, higher values shown in the Table denote a prevalence of more traditional views with respect to family and gender roles. The data highlight large differences across countries in the mean values of responses to these questions.

The share of women agreeing that men should have priority in paid work when jobs are scarce is lower among younger women than among older ones in

most countries. On average, only 12 percent of younger women agree with this statement.

The share of women believing that women need children to be fulfilled in life is also lower among younger women (with an average value of 39 percent across OECD countries) than among older women (45 percent).

On average, 80 percent of both younger and older women believe that marriage is not an outdated institution. There is almost no difference between the opinions expressed by the two cohorts of women, although the share of women agreeing with this view is higher among young women than among older ones in some countries (e.g. Belgium, Denmark, France).

The share of respondents who disapprove of women being lone parents is lower among younger women (a little over one fourth, on average) than among older ones in almost all countries.

There is much diversity in how attitudes have changed with respect to whether working mothers can have good relations with their children. On average, 17 percent of younger women consider that working mothers have worse relations to their children, a share that is marginally lower than among older women.

While a majority of women still regard being a housewife as being as satisfying as having a paid job, the proportion of young women agreeing with this statement is 3 points lower than among older women. There are wide cross-country differences in responses.

Overall, the Table points to a mixed picture, with little difference between young and older women on average with respect to the institution of marriage and the status of being a housewife (as opposed to a paid worker), but larger regarding lone parenthood, the need of children for women to be fulfilled and the presence of women in the formal labour market.

W. O.

### Reference

D'Addio, A.C. and M. M. d'Ercole (2005), “Trends and Determinants of Fertility Rates: The Role of Policies”, *OECD Social, Employment and Migration Working Papers* no. 27.

Values of women of different ages with respect to gender and family roles in percent, 1999–2001

Country	Women aged	When jobs are scarce men should have the priority to work	Difference between younger and older women	Women need to have children	Difference between younger and older women	Marriage is not an outdated institution	Difference between younger and older women	Disapprove women as lone parents	Difference between younger and older women	Working mothers cannot have a warm and stable relation with children	Difference between younger and older women	Being a housewife is as fulfilling as working in a paid job	Difference between younger and older women
Austria	15 to 34 35 to 50	14 20	-6	26 30	-3	75 78	-3	15 24	-9				
Belgium	15 to 34 35 to 50	12 24	-12	25 22	4	76 64	12	24 24	-1	14 16	-2	55 51	4
Canada	15 to 34 35 to 50	7 12	-6	14 16	-3	77 79	-2	31 34	-3	18 18	0	77 76	1
Czech Republic	15 to 34 35 to 50	13 14	-1	35 41	-6	80 91	-11	18 24	-7	21 12	8	68 72	-4
Denmark	15 to 34 35 to 50	1 3	-2	1 7	-6	87 85	2	32 34	-2	16 8	8	46 57	-10
Finland	15 to 34 35 to 50	2 5	-2	8 8	0	76 86	-11	20 19	0	5 1	4	78 77	1
France	15 to 34 35 to 50	13 21	-8	60 66	-6	67 61	6	19 26	-7	17 17	0	53 62	-9
Germany	15 to 34 35 to 50	16 18	-3	40 54	-14	74 81	-7	19 25	-5	21 34	-13	31 41	-10
Greece	15 to 34 35 to 50	8 17	-10	73 81	-8	83 84	-1	26 45	-19	17 18	-1	24 40	-16
Hungary	15 to 34 35 to 50	21 21	1	93 97	-4	83 85	-1	24 30	-6	18 26	-8	55 51	4
Iceland	15 to 34 35 to 50	2 3	0	26 29	-3	91 96	-5	11 8	3	10 4	6	59 66	-7
Ireland	15 to 34 35 to 50	2 14	-12	4 11	-6	69 78	-9	18 28	-11				
Italy	15 to 34 35 to 50	12 22	-10	46 56	-10	81 85	-5	31 34	-3	25 25	0	39 49	-10
Japan	15 to 34 35 to 50	16 27	-11	39 53	-14	86 93	-7	14 24	-10	3 5	-1	81 86	-5



(Table continued)

Country	Women aged	When jobs are scarce men should have the priority to work	Difference between younger and older women	Women need to have children	Difference between younger and older women	Marriage is not an outdated institution	Difference between younger and older women	Disapprove women as lone parents	Difference between younger and older women	Working mothers cannot have a warm and stable relation with children	Difference between younger and older women	Being a housewife is as fulfilling as working in a paid job	Difference between younger and older women
Luxembourg	15 to 34 35 to 50	12 18	-7	31 32	-1	68 66	3	25 28	-3	15 19	-4	54 57	-3
Mexico	15 to 34 35 to 50	26 31	-5	37 48	-11	82 77	5	34 41	-8	28 26	2	70 71	-1
Netherlands	15 to 34 35 to 50	2 11	-9	2 3	-1	78 71	7	19 14	5	12 14	-3	37 45	-8
Poland	15 to 34 35 to 50	19 29	-10	51 66	-15	91 91	-1	12 16	-4	34 38	-4	54 47	7
Portugal	15 to 34 35 to 50	21 25	-4	57 75	-18	75 70	5	38 44	-6	22 36	-14	39 44	-4
Korea	15 to 34 35 to 50	14 44	-30	82 93	-10	75 84	-9	61 84	-24	9 15	-6	95 93	2
Slovak Republic	15 to 34 35 to 50	12 14	-2	42 45	-3	88 92	-4	23 26	-3	12 18	-6	64 63	1
Spain	15 to 34 35 to 50	10 14	-4	35 43	-9	74 81	-7	5 10	-4	14 14	1	46 52	-5
Sweden	15 to 34 35 to 50	1 0	0	25 21	4	88 71	16	36 35	1	11 12	-1	47 50	-3
Turkey	15 to 34 35 to 50	49 58	-9	71 81	-10	92 92	0	87 92	-4	26 31	-4	73 77	-4
United Kingdom	15 to 34 35 to 50	9 16	-8	14 19	-4	67 78	-11	21 26	-5	21 21	0	58 56	2
United States	15 to 34 35 to 50	8 7	1	10 16	-5	89 91	-1	36 42	-6	13 18	-4	76 76	-1
OECD-26	15 to 34 35 to 50	12 19	-6	39 45	-6	80 81	-1	27 32	-5	17 19	-2	58 61	-3

Source: Calculations based on data extracted from different waves of the World Values Survey.

## TRENDS IN DEVELOPMENT AID (2)

The size, regional and sectoral structure of development assistance as well as its change over time was investigated in a previous article on this subject (DICE Report 2005). The focus of this article is on tied aid, debt relief and donor coordination. The structures of development administration will be treated in the next issue of the DICE Report. All three articles refer to the discussion of the UN “Millennium Development Goals”, to be attained in 2015.

### Tied aid

Tying development aid to exports of the donor country played an important role in the early years of post-World War II development assistance. Tying aid eliminates – or at least reduces – the possibility of the recipient country freely choosing its imports and selecting the lowest price supplier. Tied aid is, thus, of lower value to the recipient than the face value of the aid amount suggests. Today, tied aid has significantly lost importance. On average, for the 20 donors considered here, only about 15 percent of the aid in 2003 was still tied to donors’ exports (Table, col. 1). However, there are considerable differences between donors. Some countries, such as Spain or Sweden, abstain completely from tying, while Belgium, Finland or Portugal still tie a considerable share of assistance to their own exports. Though the importance of tying is reduced, one has to bear in mind that the tying percentage relates to exports of goods and some services, but, for example, not to the so-called technical assistance, which – contrary to what one might understand with this term – is consultancy. Such consultancy is typically rendered by experts who are nationals of the donor country, although often similar advice can also be provided at a lower cost, by experts from India, for example.

### Debt relief

A part of development assistance is given as credit not as grants. To qualify as development assistance, according to the criteria of the Development Assistance Committee (DAC) of the OECD countries, the credit must entail a grant element of at least 25 percent, i.e., the interest rate must be at least 25 percent below the market interest rate, while the DAC target

norm for development credits is a grant element of more than 80 percent. Any credit with a grant element of less than 100 percent must be served and finally paid back. Despite considerable grant elements included in the development credits, specifically in credits to very poor countries, many recipients of aid have repeatedly gone into debt serving difficulties and even debt crises, which made any hope for a final, orderly settlement of the debt futile – not least with respect to the political and social stability in the recipient country. In such cases the donor countries (those within the “Paris Club”) have already agreed several times to cancel parts of the outstanding debt, accumulated from continued public development credits. (The “London Club” does the same for problems with private sector loans.) However, for many recipient countries the underlying problem cannot be solved by repeated (partial) cancellation of debt – in fact, the repetition and anticipation of debt cancellations may even be a part of the problem.

In 2003 again, a part of outstanding debt has been cancelled (col. 3 and 4 in the Table), particularly those of the poorest, the “least developed” countries. However, one should not think that debt cancellation does not affect the current flows of development aid. In fact, it might reduce the flow of fresh money, because the amount of debt cancelled is counted as development aid (more precisely as gross aid, not as net aid) which means that the current aid payments can be reduced without increasing the distance of the donor and debt relieving country to the 0.7 percent norm (DICE Report 2005). In 2003 the 20 donor countries considered here cancelled debts of US\$9.2 billion or 14 percent of total 2003 net aid.

The bulk of that amount (\$7.2 billion, not shown in the Table) benefited the poorest countries. Belgium, France, Italy and Germany showed the strongest commitment to debt relief for these countries. In general, nearly all analyzed donor countries show a clear increase in bilateral debt relief compared to the figures of 1998 and 1993 (likewise not shown).

### Donor coordination and policy coherence

Donors have long been criticized for a lack of coordination in three areas: coordination with the recipients’ national policy goals, coordination with other donors, and coordination within their own political institutions, many of which are involved in the administration of donor countries’ development assis-

Typing of aid, debt relief, bilateral and multilateral aid, 2003

	Typing in % of total ODA	ODA net in million US\$	Debt relief in million US\$	Debt relief in % of total ODA	Total gross		Multilateral ODA in % of bilateral ODA
					bilateral ODA in million US\$	multilateral ODA in million US\$	
Austria	6.28	505	39	7.7	228	276	121.0
Belgium	48.65	1,853	753	40.6	1,468	385	26.0
Denmark	0.98	1,748	10	0.6	1,032	717	69.0
Finland	86.23	558	6	1.2	309	250	81.0
France	14.16	7,253	2,936	40.5	5,213	2,040	39.0
Germany	3.06	6,784	1,284	18.9	4,060	2,724	67.0
Greece	5.39	362	2	0.6	228	134	58.0
Ireland	0	504	3	0.6	352	152	43.0
Italy	n.a.	2,433	598	24.6	1,061	1,372	129.0
Luxembourg	n.a.	194	–	–	150	44	30.0
Netherlands	n.a.	3,981	257	6.4	2,951	1,030	35.0
Portugal	43.99	320	6	2.0	182	137	75.0
Spain	0	1,961	106	5.4	1,151	810	70.0
Sweden	0	2,400	165	6.9	1,779	621	35.0
United Kingdom	4.94	6,282	161	2.6	3,861	2,421	63.0
European Communities	n.a.	n.a.	n.a.	–	6,445	728	11.0
Norway	0.13	2,042	69	3.4	1,462	580	40.0
Switzerland	3.58	1,299	75	5.8	945	355	38.0
Japan	3.4	8,880	940	10.6	6,334	2,545	40.0
USA	n.a.	16,254	1,800	11.1	14,594	1,661	11.0
Average	15.3	3,453	512	10.7	2,690	949	–
Sum	–	65,613	9,210	14.0	53,805	18,982	35.3

Source: OECD, DAC online ([www.oecd.org/dac/stats/idsonline](http://www.oecd.org/dac/stats/idsonline)).

tance. This lack of coordination, which can lead to an incoherent development aid policy, is also held responsible for the apparently low effect of development aid and for the developmental standstill in many poor countries.

One mechanism to improve aid coordination is the delivery of assistance via multilateral institutions such as the World Bank and different UN bodies. For European donor countries, an obvious candidate for multilateral lending is the European Commission, which has gained increasing importance in recent years. In 2003, it received by far the biggest share of the total funds dedicated to multilateral institutions by the fifteen EU donor countries, with Germany, France and the UK being the three main contributors. Col. 5 of the Table shows that \$6,445 million was spent on bilateral aid and \$728 million on multilateral aid by the EU. Both amounts stem mostly from multilateral contributions made by EU member countries. The Commission itself uses the bulk of that amount for bilateral assistance while a small part (\$728 million) is forwarded to (other) multilateral donors like the World Bank.

Germany also heads the list with regard to the individual donor's overall multilateral disbursements with an amount of \$2,724 million in 2003, a sum which equals two thirds of Germany's bilateral disburse-

ments for the same year. In comparison, the United States only disbursed an amount equal to 11 percent of their bilateral aid to multilateral agencies.

On average, multilateral aid is only slightly more than a third of bilateral aid in the 20 researched countries (including aid from the European Commission). Therefore, the bulk of development assistance is provided without the potentially coordinating effect of multilateral donor institutions. Arguments in favour of such coordinating mechanisms are not only based on the motivation of exploiting maximal synergies between the different donors' development efforts but also in hope of reducing the recipient countries' management costs caused by foreign aid. Obviously, however, the direct contact to and influence on the recipient countries is what the donor countries want to have and try to exert via bilateral assistance.

Critics of multilateral aid point to the immense administrative costs some multilateral agencies have to finance via the through-flow of aid. These critics argue instead for a better coordination of bilateral donors. To support coordination efforts, the DAC, itself an important forum for donor coordination, has developed a set of "Principles for Country Ownership, Donor Harmonization and Alignment" (OECD DAC, 2005). The main points are that the donors respect and support the development priorities of the recipient coun-

tries, that they coordinate their developmental efforts under the guidance of the recipient country and that they strengthen the capability of the recipient countries to plan and implement development projects.

There are signs that the donor community is taking aid coordination more seriously than in the past. An example is Mozambique. Some years ago the country introduced comprehensive reforms of the institutional setup of its economy and economic policy, which were supported by the World Bank. After a promising start of the reforms 16 DAC members decided to contribute to a “Joint Program for Macroeconomic Support”, which in effect provides funds for the country’s national budget. This procedure is – on one hand – a maximum of coordination of bilateral donors, but – on the other – also a return to “budget aid”, which had been largely abandoned in favour of project aid many years ago.

R. O. and A. R.\*

## References

CESifo DICE Report (2005), “Trends in Development Aid”, vol. 3, no. 3, 66–69.

Dollar, D. and V. Levin (2004), “The Increasing Selectivity of Foreign Aid 1984–2002”, *Policy Research Working Paper*, no. 3299, Policy Research Group, World Bank.

Erixon, F. (2005), “Aid and Development. Will it Work this Time?”, *International Policy Network*.

Harford, T. and M. Klein (2004), “Donor Performance. What do we know and what should we know?”, *Public Policy for the Private Sector*, note 278, Washington DC.

OECD DAC (2005), *The DAC Journal of Development Co-operation*, vol. 6, no. 1.

World Bank (2001), *World Development Indicators 2000*, Washington DC.

---

\* Anja Rucker is development consultant to the South African government. She works on behalf of the German Agency for Technical Cooperation (GTZ).

## WHERE IS DOING BUSINESS EASY?

In “Doing Business in 2006” the World Bank investigated the regulations that enhance business activity and those that constrain it. As a result ten quantitative indicators with respect to business regulations and their enforcement can be compared across 155 countries. There were three more indicators this year than in the last investigation “Doing Business in 2005”: dealing with licences, paying taxes and trading across borders.

The new added indicator “dealing with licenses” analyses procedures, time and cost of business inspections and licensing in the construction industry. Most reforms of building codes are driven by concerns about safety. Stricter codes result in fewer deaths – except when regulation is so burdensome that construction moves to the informal sector. In New Zealand a builder would have to complete

7 procedures requiring 65 days and \$6,800 – a third of the average annual income – to comply with all regulations. That excludes the time and cost of building. Inspections are contracted to private companies. The efficiency of regulatory and inspection services makes New Zealand the world’s easiest place to build a warehouse. In Denmark it takes 7 procedures, 70 days and \$ 31,800. In Canada it is no longer necessary for a builder to visit municipalities again and again to receive lists with technical requirements to build a warehouse. Regulations issued in 2004 limit municipalities to a single list of requirements. Repeat visits are not necessary. It now takes less than 3 months to fulfil the requirements for building a warehouse in Toronto, putting Canada among the fastest places to complete the process (Table 1).

The second new indicator in this year’s Doing Business is “paying taxes”. Simple, moderate taxes and fast, cheap administration mean less hassle for business – as well as higher revenues. Burdensome taxes create incentives for evasion. Businesses prefer low-

**Table 1**

### Dealing with licenses <sup>a)</sup>

	Proce- dures number	Time days	Cost in % of income per capita
Austria	14	195	81.6
Belgium	15	184	64.1
Czech Republic	31	245	16.1
Denmark	7	70	71.3
Estonia	12	116	41.4
Finland	17	56	76.2
France	10	185	78.3
Germany	11	165	82.8
Greece	17	176	71.9
Hungary	25	213	279.1
Ireland	10	181	23.6
Italy	17	284	147.3
Latvia	21	160	43.9
Lithuania	14	151	17.5
Netherlands	18	184	142.7
Poland	25	322	83.1
Portugal	20	327	57.7
Slovak Republic	13	272	18.0
Slovenia	14	207	128.7
Spain	12	277	77.1
Sweden	8	116	119.6
United Kingdom	19	115	70.2
Norway	13	97	53.9
Switzerland	15	152	59.2
Australia	16	121	12.3
Canada	15	87	123.0
Japan	11	87	19.7
New Zealand	7	65	29.3
United States	19	70	16.9

<sup>a)</sup> For a business in the construction industry to build a standardized warehouse.

Source: World Bank, Doing Business 2006, database.

**Table 2**

### Paying taxes

	Pay- ments number	Time hours	Total tax payable in % of gross profit
Austria	20	272	50.8
Belgium	10	160	44.6
Czech Republic	14	930	40.1
Denmark	18	135	63.4
Estonia	11	104	39.5
Finland	19	n.a.	52.1
France	29	72	42.8
Germany	32	105	50.3
Greece	32	204	47.9
Hungary	24	304	56.8
Ireland	8	76	45.3
Italy	20	360	59.8
Latvia	39	320	38.7
Lithuania	13	162	41.6
Netherlands	22	700	53.3
Poland	43	175	55.6
Portugal	7	328	45.4
Slovak Republic	31	344	39.5
Slovenia	29	272	47.3
Spain	7	56	48.4
Sweden	5	122	52.6
United Kingdom	22	n.a.	52.9
Norway	3	87	60.1
Switzerland	25	63	22.0
Australia	12	107	37.0
Canada	10	119	32.5
Japan	26	315	34.6
New Zealand	8	70	44.2
United States	9	325	21.5

Source: World Bank, Doing Business 2006, database.



er tax rates that are simply applied. Or, if rates are high, businesses want something in return for tax payments. A complicated tax system costs a lot of money to run – funds that could be better spent on education, health care or infrastructure. A smaller tax burden encourages firms to invest. Lower rates work best when their administration is simple. The World Bank has analysed the question “who makes paying taxes easy – and who does not?”

Considering the number of annual payments of taxes Hong Kong ranks first in the world with just one payment. With three payments p.a. Norway is one of the countries worldwide with the lowest number of tax payments. Of the additional OECD countries, Sweden (5 payments p.a.), Portugal, Spain, Ireland and New Zealand with seven and eight payments, respectively, are among the top ten countries. With 43 Poland has the highest number of payments among the OECD countries.

Worldwide the time needed to complete the forms is least in the United Arab Emirates with 12 hours annually. In the OECD countries it seems to be much more difficult to complete the tax files: only Spain, Switzerland and New Zealand are among the ten countries where filing involves less time. People in the Czech Republic need the most time to comply with tax requirements (930 hours p.a.). This is one of the most time-consuming filing procedures worldwide.

The tax rates for companies range within the OECD countries from 21.5 percent and 22.0 percent in the United States and Switzerland, respectively, to 63.4 percent of gross profit in Denmark (Table 2).

The last new indicator measured by the World Bank is trading across borders. “Doing Business” has compiled the procedural requirements for exporting and importing goods. The time is counted from the time

the business starts preparing the necessary documents to the time the cargo is in the client’s warehouse (excluding the days on board). Every official procedure is counted – from the contractual agreement between the two parties to the delivery of goods – along with the time necessary for completion. All documents and signatures required for clearance of the goods across the border are also recorded. The fewest number of days to export goods is needed by exporters from Denmark, Germany, Lithuania and Sweden (between 5 and 6). Among the OECD countries the most days needed by exporters are from Greece (29). Only one signature for clearance of goods is needed in Germany and Sweden. The exporters from Australia, Austria, Canada, Denmark, Estonia and New Zealand just need two. The most signatures are needed in the Slovak Republic (9).

Table 3

## Trading across borders

	Documents for export number <sup>a)</sup>	Signatures for export number <sup>b)</sup>	Time for export days <sup>c)</sup>	Documents for import number <sup>a)</sup>	Signatures for import number <sup>b)</sup>	Time for import days <sup>c)</sup>
Austria	4	2	8	5	3	9
Belgium	5	2	7	6	2	9
Czech Republic	5	3	20	8	4	22
Denmark	3	2	5	3	1	5
Estonia	5	2	12	5	5	14
Finland	4	3	7	3	1	7
France	7	3	22	13	3	23
Germany	4	1	6	4	1	6
Greece	7	6	29	11	9	34
Hungary	6	4	23	10	5	24
Ireland	5	5	14	4	5	15
Italy	8	5	28	16	10	38
Latvia	9	6	18	13	7	21
Lithuania	5	5	6	12	4	17
Netherlands	5	3	7	4	1	8
Poland	6	5	19	7	8	26
Portugal	6	4	18	7	5	18
Slovak Republic	9	8	20	8	10	21
Slovenia	9	7	20	11	9	24
Spain	4	3	9	5	3	10
Sweden	4	1	6	3	1	6
United Kingdom	5	5	16	4	5	16
Norway	4	3	7	4	3	7
Switzerland	8	5	21	13	5	22
Australia	5	2	12	11	2	16
Canada	6	2	12	7	1	12
Japan	5	3	11	7	3	11
New Zealand	5	2	8	9	2	13
United States	6	5	9	5	4	9

Note: The main indicators for trading across borders measure procedural requirements for exporting and importing a standardized cargo of goods. They include: <sup>a)</sup> Number of all documents required to export/import goods. – <sup>b)</sup> Number of all approvals, signatures or stamps that are required to export/import goods. – <sup>c)</sup> Time necessary to comply with all procedures required to export/import goods.

Source: World Bank, Doing Business 2006, database.

Among the OECD countries it takes importers in Italy the most days – 38 – to import goods. The countries where exporting is the easiest, importing is the also easiest.

All in all trading across borders is easiest in Denmark, followed by Sweden; Germany is third, then Finland, the Netherlands, Norway, Austria, Belgium and then Spain (Table 3).

But rankings on the ease of doing business do not tell the whole story. The indicator is limited in scope. It does not account for a country's proximity to large markets, its quality of infrastructure services, security of property from theft and looting, macroeconomic conditions or the underlying strength of institutions.

N. H.

### Reference

World Bank (2005), *Doing Business in 2006: Creating Jobs*, Washington, DC.

## MORTGAGE DEBT

Mortgage debt in percent of GDP differs widely across the EU-15 countries (Table 1). The Netherlands, Denmark and the UK exhibit percentages between 70 and nearly 100, while Italy, Greece and France are in the range between 13 and 25 percent. The share in the European mortgage market is highest for Germany and the UK with the Netherlands ranking third. The real mortgage growth rate (col. 3 of Table 1), too, differs considerably between countries, but is in almost all cases significantly higher than the GDP growth rate. Mortgage growth seems to be high in countries with a low amount of mortgage debt per GDP (like Italy and Greece), while it is low in countries with an already high percentage of mortgage debt (like UK and Germany).

The figures of Table 1 may be partly explained by the different tax treatment for housing and mortgages (Table 2). But, as Wolswijk in a recent analysis (2005) shows, more important determinants of mortgage credit are the stock market growth, increases of the price for houses and financial deregulation.

Developments on the mortgage market are relevant for macroeconomic policy decisions because interest rate setting for mortgages may impact on the monetary policy transmission channel. Moreover, “mortgage debt” seems not to be limited to housing purposes

**Tabelle 1**  
**EU mortgage and housing market characteristics, 2003**

	Mortgage debt (in % of GDP)	Share in EU mortgage market in %	Real mortgage growth rate in %
Austria	26.4	1.4	9.0
Belgium	28.5	1.8	3.5
Denmark	87.5	3.9	4.8
Finland	35.6	1.2	7.6
France	24.7	9.1	7.7
Germany	54.3	27.2	0.9
Greece	17.4	0.6	21.6
Ireland	45.0	1.4	21.7
Italy	13.3	4.1	18.5
Luxembourg	33.4	0.2	15.6
Netherlands	99.9	10.7	14.3
Portugal	50.6	1.6	-1.1
Spain	42.1	7.4	16.4
Sweden	50.0	3.2	5.7
United Kingdom	70.4	26.4	3.4

Source: Wolswijk (2005) and sources mentioned there.

**Tabelle 2**  
**Main tax categories affecting housing/mortgages in EU, 2003**

	Tax on imputed rent	Interest deductibility	Capital gains tax
Austria	N	Y	n
Belgium	Y	Y	n
Denmark	N	Y	n
Finland	N	Y	n
France	N	N	n
Germany	N	N	n
Greece	N	Y	N
Ireland	N	Y	n
Italy	Y	Y	n
Luxembourg	Y	Y	n
Netherlands	Y	Y	N
Portugal	N	Y	n
Spain	N	Y	n
Sweden	N	Y	n
United Kingdom	N	N	n

Note for the last column: Capital letters denote absence of a capital gains tax, small letters refer to a capital gains tax being in place, but the gains on selling a house being de facto exempted, for instance because it is sold a certain number of year after its acquisition.

Source: International Bureau of Fiscal Documentation (2003), compiled by Wolswijk (2005).

es proper, but, in practice, applies also to non-housing consumption purposes. Finally, different housing and mortgage tax treatments across countries and, thus, different degrees of housing subsidies may increase the differences between national inflation rates (van den Noord 2003, Osterkamp 2005).

This leads to the question whether the use of fiscal instruments to correct imbalances and price volatility on the housing market and to address distortions of the monetary policy transmission channel is in order. The following instruments – specifically with view to price volatility – are analysed by Wolswijk: increased stamp duties (transaction tax on buying a house), regular updating of market values of dwellings, inclusion of anti-cyclical (or anti-speculative) elements in the capital gains tax, targeted reductions in mortgage interest tax relief (e.g. for low interest rates). Moreover, municipalities could increase the elasticity of housing supply by adapting the release of building permits to the situation on the housing market.

Such instruments are considered by Wolswijk to be of a structural nature. But he also asks whether “fine-tuning” may be warranted and possible. The author mentions changes of or exemptions from the turnover tax for building and repairing houses, cycle-dependen-

dent changes of stamp duties, property tax rates and subsidies. He mentions an earlier idea of Maclennan et al. who propose for the UK to make the Bank of England responsible for setting the property tax rate. Wolswijk is, however, cautious what concerns the use of fine-tuning instruments due to doubts about their appropriate timing and calibration.

R. O.

## References

Maclennan, D., J. Muellbauer and M. Stephens (2003), "Asymmetries in Housing and Financial Market Institutions and EMU", *CEPR Discussion Paper* 2062.

Noord, P. van den (2003), "Tax Incentives and House Price Volatility in the Euro Area: Theory and Evidence", *OECD Economics Department Working Paper* 356.

Osterkamp, R. (2005), "Tax Preferences for Housing", *DICE Report* 3(1), 83 – 84.

Osterkamp, R. (2004), "Mortgage Banks", *DICE Report* 2(4), 67– 71.

Wolswijk, G. (2005), "On some Fiscal Effects on Mortgage Debt Growth in the EU", *European Central Bank Working Paper* 526.

## NATIONAL COLLECTIVE BARGAINING SYSTEMS IN EUROPE

Collective bargaining plays a key role in industrial relations in all current EU member states, though national systems differ widely in terms of the level, coverage, content and nature of bargaining. Despite all these differences, industrial relations research usually distinguishes between two groups of countries that have a broad range of similarities in their collective bargaining systems. The first group, which covers most of the “old” EU member states, is characterised by relatively strong multiemployer bargaining institutions, with sectoral or intersectoral bargaining and relatively high bargaining coverage. This rather strong and centralised system of collective bargaining differs fundamentally from bargaining systems in most countries outside Europe as, for example, in the US or Japan. There is, however, a second group within Europe that includes the UK and most of the new EU member states, which has comparatively weak bargaining institutions with the company as the predominant bargaining level and relatively low bargaining coverage.

Beyond this relatively clear-cut dichotomy between centralised and decentralised bargaining systems in practice many European countries have a quite differentiated multi-level bargaining system in which different levels of bargaining (intersectoral, sectoral and company) are connected with each other. Moreover, other forms of negotiations such as tripartite consultations at national level or “works agreements” at company level (i.e. concluded by works councils or similar bodies) are often closely connected with collective bargaining.

### *Main characteristics of national collective bargaining systems*

There are several elements that define national collective bargaining systems. The first is the relative importance of the various bargaining levels, with a differentiation between more centralised and more decentralised bargaining systems. A second element is bargaining coverage, which is influenced by the levels of bargaining as well as by the existence and use of extension procedures. A third element is whether or not tripartite concertation at the national level has an influence on collective bargaining.

Since many European countries have multi-level bargaining systems, the relationships and interactions between the different bargaining levels can be treated as a fourth main element characteristic of a national bargaining system.

### *Levels of bargaining*

The Table outlines the importance of the various levels of bargaining in the 26 countries considered. The importance of the different bargaining levels may vary depending on the topics regulated by collective agreements and, therefore, the information provided in the Table focuses exclusively on wage bargaining. In this respect, three groups of countries may be distinguished:

- The first and smallest group covers four countries (Belgium, Finland, Ireland and Slovenia) where the intersectoral level is the most important bargaining level for the determination of wages. In addition, there are five countries (Greece, Estonia, Hungary, Lithuania and Romania) where the national minimum wage is determined by bipartite or tripartite concertation at the national level. Since the coverage of the other bargaining levels (sectoral and company level) is much lower in these countries, one might consider the intersectoral level the most important bargaining level. However, except for the national minimum wage, lower bargaining levels are most important for the determination of wages in these countries;
- A second group of 11 countries, mostly in northern and western Europe (Austria, Bulgaria, Denmark, Germany, Greece, Italy, the Netherlands, Norway, Spain, Slovakia and Sweden), have national bargaining systems in which the sector is the most important bargaining level for wage determination;
- A third group of 10 countries, including most of the CEE states (the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland and Romania) plus Cyprus, Malta and the UK, have relatively decentralised bargaining systems with company bargaining dominant.

Finally, there is France, which does not fit into any of these groups since it has no pay bargaining level that is clearly most important. While in particular for small and medium-sized companies the sector level is the most important, it is the company level that is key for most large companies.



**Levels of collective wage bargaining, bargaining coverage and influence of tripartite concertation  
(latest available figures)**

	Importance of bargaining level			Collective bargaining coverage in %	Influence of tripartite concertation	Tripartite consultation on national minimum wage
	Inter sectoral level	Sectoral level	Company level			
<b>Intersectoral bargaining dominant</b>						
Belgium <sup>a)</sup>	XXX	X	X	> 90	Yes	Yes
Finland	XXX	X	X	+/-90	Yes	No
Ireland <sup>b)</sup>	XXX	X	X	> 44	Yes	Yes
Slovenia	XXX	X	X	< 100	Yes	Yes
<b>Sectoral bargaining dominant</b>						
Austria		XXX	X	98-99	No	No
Bulgaria	X	XXX	X	25-30	No	Yes
Denmark <sup>c)</sup>	X	XXX	X	+/-77	No	No
Germany <sup>d)</sup>		XXX	X	+/-70	No	No
Greece	(XXX)	XXX	X	60-70	No	No
Italy		XXX	X	+/-90	Yes	No
Netherlands	X	XXX	X	+/-80	Yes	No
Norway	XX	XXX	X	70-77	Yes	No
Spain	X	XXX	X	+/-80	Yes	Yes
Slovakia		XXX	X	+/-40	No	Yes
Sweden		XXX	X	> 90	No	No
<b>No bargaining level clearly dominant</b>						
France	X	XX	XX	+/-90	No	Yes
<b>Company bargaining dominant</b>						
Cyprus <sup>e)</sup>		X	XXX	27	No	No
Czech Republic		X	XXX	25-30	Yes	Yes
Estonia	(XXX)	X	XXX	20-30	Yes	Yes
Hungary	(XXX)	X	XXX	+/-40	Yes	Yes
Latvia		X	XXX	10-20	Yes	Yes
Lithuania	(XXX)	X	XXX	+/-10	Yes	Yes
Malta <sup>f)</sup>		X	XXX	+/-50	Yes	Yes
Poland		X	XXX	+/-40	Yes	Yes
Romania <sup>g)</sup>	(XXX)	X	XXX	n.a.	Yes	Yes
United Kingdom		X	XXX	< 40	No	Yes

X = existing level of wage bargaining; XX = important, but not dominant level of wage bargaining; XXX = dominant level of wage bargaining; (XXX) = bargaining on national minimum wage.

<sup>a)</sup> Consultation on the minimum wage in the sense that the social partners will probably consult the government if they plan to modify the minimum wage. – <sup>b)</sup> There are no figures on Irish bargaining coverage available, but coverage must be above 44.5 percent (which is the union density rate) since all union members are automatically covered by national agreements, while many non-union employees de facto receive the nationally agreed pay increases. – <sup>c)</sup> There is one main intersectoral agreement covering all manufacturing sectors in Denmark; bargaining coverage refers to private sector only (it is almost 100 percent in the public sector). – <sup>d)</sup> Bargaining coverage refers to West Germany – in East Germany bargaining coverage is only about 54 percent. – <sup>e)</sup> Bargaining coverage refers to Cypriot private sector only (it is almost 100 percent in the public sector). – <sup>f)</sup> There is automatic annual adjustment of wages to price developments in Malta; different studies estimate the proportion of employees covered by collective agreements at between 40 percent and 60 percent. – <sup>g)</sup> All employees in Romania are covered by the national agreement on minimum wages; no figures are available on the coverage of sectoral and company agreements, but it is estimated that a large proportion of employees are not covered by these agreements.

Sources: EIRO; European Commission, Industrial Relations in Europe 2004.

### Bargaining coverage

There are huge differences in collective bargaining coverage across Europe. The proportion of employees covered by collective agreements varies between nearly 100 percent in Austria or Slovenia to around 10 percent in Lithuania. There are two main ele-

ments that influence the bargaining coverage. First, there is a clear correlation between bargaining coverage and the relative importance of bargaining levels. Countries with a dominance of sectoral or intersectoral bargaining tend to have much higher bargaining coverage than countries where company bargaining predominates. Second, the existence and use

of extension procedures may have a positive effect on bargaining coverage. This is the case, for example, in France, Italy and the Netherlands. However, there are many countries, in particular in Central and Eastern Europe, which have introduced the possibility of an extension procedure but have not used it very often so far, and bargaining coverage is still rather low.

#### *Influence of tripartite concertation*

In 16 out of the 26 countries investigated, tripartite concertation at the national level has a regular influence on wage bargaining. In three countries (Finland, Ireland and Slovenia) the state sits directly at the bargaining table so that collective agreements at the intersectoral level are concluded on a tripartite basis. In many other countries (e.g. the Netherlands, Norway, Spain and many of the CEE countries) there are either regular or irregular tripartite framework agreements at the national level that contain legally non-binding recommendations for pay increases. In addition, almost all countries with a statutory minimum wage have also some form of tripartite consultation on its regular adjustment. This also applies to countries such as France or the UK that have no other form of formal tripartite concertation.

W. O.

#### **Reference**

EIRO (2005), *Changes in National Collective Bargaining Systems since 1990*, Dublin (eironline).

## TRADE RESTRICTIONS

What is the present overall level of restrictions to international trade and how has this level developed over time? This question is only simple on the surface. A satisfactory answer is difficult to achieve although the effects such an answer could have for economic policy making are far reaching.

The basic difficulty of arriving at an overall measurement of trade restrictions lies in the task of aggregation. One problem is to make a vast variety of different types of non-tariff barriers commensurable among each other as well as with (the easier case of) tariff barriers. Another problem is to do this for virtually innumerable trade flows between all countries and for all types of goods.

In the 1980s and 1990s important steps at a conceptual and empirical level for developing an overall measure for trade restrictions were taken, for example by the World Bank (Laird and Yeats 1988), the Global Trade Analysis Project (GTAP, Gehlhar 1996), the OECD (1997a, 1997b) and Coppel and Durand (1999). More recent initiatives, again with respect to theory and empirics, are found in the research of the World Bank (2004, 2005), the OECD (2004), UNCTAD (2005) and the IMF (2005). Single researchers too, such as Busse (2002) or Anderson and Neary (2004), have contributed to a deeper understanding and to a solution of the problem: Busse with a discussion paper on a special aspect of trade restrictions; Anderson and Neary with a book treating the subject comprehensively.

The empirical results of the different research approaches are seldom directly comparable, due to different delimitations of groups of countries, different aggregation levels of goods and different theoretical concepts and kinds of measurement. Moreover, the meas-

ured trade restriction applied by an individual country is mostly expressed in one figure, which does not differentiate between the different trade restrictions vis-à-vis its individual trade partners. Thus, it is possible to relate the amount and development of imports of a country with regard to the trade restrictions the country imposes. But it is, unfortunately, not possible to look at the more interesting question, namely how exports of the countries affected by restrictions react to the restrictions they face.

Fortunately, there is one exception. Kee, Nicita and Olarreaga (all at the World Bank, 2005) have calculated the trade restrictions *between* important groups of participants in world trade. Their empirical results are presented in the Table. OTRI stands for the authors' Overall Trade Restrictiveness Index. The value of the index is the answer to the question: What is

Figure 1

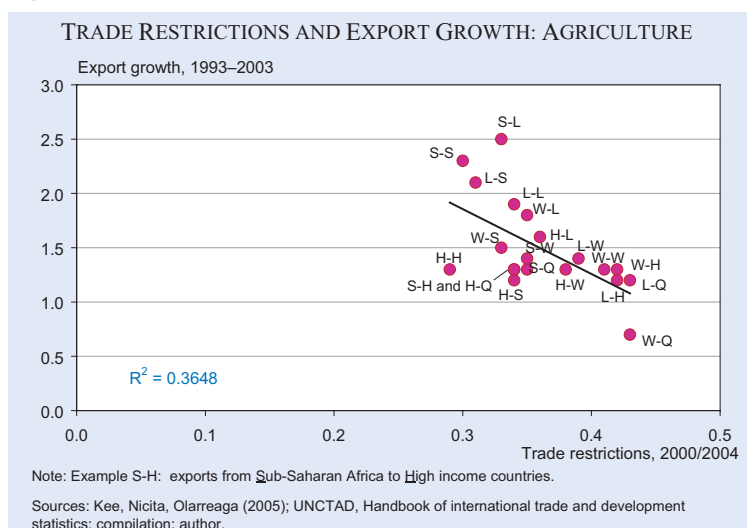
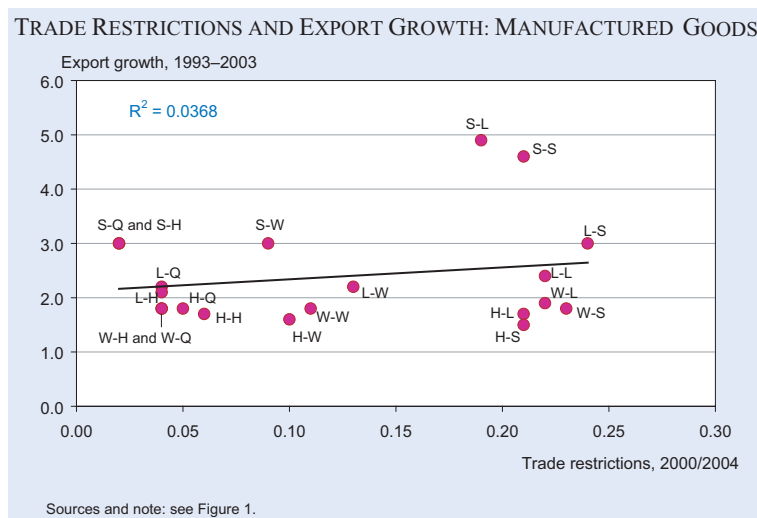


Figure 2



## Overall bilateral trade restrictions between seven groups of countries

Exporters	Importers						
	QUAD	High Income	Middle Income	Low Income	LDC	SSA	World
<b>A. Overall Trade Restrictiveness Index: Agriculture and manufacturing</b>							
QUAD	0.08	0.08	0.19	0.23	0.21	0.23	0.14
<b>High Income</b>	<b>0.08</b>	<b>0.08</b>	0.19	<b>0.23</b>	0.21	<b>0.23</b>	<b>0.14</b>
Middle Income	0.09	0.09	0.22	0.25	0.22	0.25	0.15
<b>Low Income</b>	<b>0.14</b>	<b>0.14</b>	0.25	<b>0.26</b>	0.22	<b>0.26</b>	<b>0.20</b>
LDC	0.12	0.12	0.24	0.25	0.22	0.26	0.18
<b>SSA</b>	<b>0.11</b>	<b>0.11</b>	0.23	<b>0.24</b>	0.21	<b>0.24</b>	<b>0.17</b>
<b>World</b>	<b>0.10</b>	<b>0.09</b>	0.22	<b>0.24</b>	0.22	<b>0.25</b>	<b>0.15</b>
<b>B. Overall Trade Restrictiveness Index: Agriculture</b>							
QUAD	0.30	0.31	0.46	0.37	0.29	0.34	0.37
<b>High Income</b>	<b>0.34</b>	<b>0.29</b>	0.46	<b>0.36</b>	0.28	<b>0.34</b>	<b>0.38</b>
Middle Income	0.49	0.47	0.42	0.36	0.28	0.33	0.43
<b>Low Income</b>	<b>0.43</b>	<b>0.42</b>	0.40	<b>0.34</b>	0.27	<b>0.31</b>	<b>0.39</b>
LDC	0.38	0.38	0.39	0.32	0.26	0.30	0.37
<b>SSA</b>	<b>0.35</b>	<b>0.34</b>	0.38	<b>0.33</b>	0.26	<b>0.30</b>	<b>0.35</b>
<b>World</b>	<b>0.43</b>	<b>0.42</b>	0.42	<b>0.35</b>	0.28	<b>0.33</b>	<b>0.41</b>
<b>C. Overall Trade Restrictiveness Index: Manufacturing</b>							
QUAD	0.06	0.06	0.15	0.21	0.20	0.21	0.11
<b>High Income</b>	<b>0.05</b>	<b>0.06</b>	0.15	<b>0.21</b>	0.20	<b>0.21</b>	<b>0.10</b>
Middle Income	0.04	0.04	0.19	0.23	0.21	0.24	0.11
<b>Low Income</b>	<b>0.04</b>	<b>0.04</b>	0.20	<b>0.22</b>	0.21	<b>0.24</b>	<b>0.13</b>
LDC	0.03	0.03	0.20	0.22	0.21	0.24	0.12
<b>SSA</b>	<b>0.02</b>	<b>0.02</b>	0.16	<b>0.19</b>	0.18	<b>0.21</b>	<b>0.09</b>
<b>World</b>	<b>0.04</b>	<b>0.04</b>	0.18	<b>0.22</b>	0.20	<b>0.23</b>	<b>0.11</b>
Notes: The higher the figure, the stronger is the trade restriction. The figures in <b>bold</b> indicate those trade relations for which we could find export figures.							

Source: Kee, Nicita, Olarreaga (2005).

the equivalent uniform tariff of a country that would keep imports of that country at the observed level? The figures relate to the years between 2000 and 2004. QUAD indicates the four main players of world trade, namely the EU, US, Japan and Canada. The QUAD countries form a sub-group of the high income countries, while the countries of Sub-Saharan Africa (SSA) are a sub-group of the low income countries.

A complementing fortunate circumstance is that UNCTAD, in its *Handbook of International Trade and Development Statistics*, publishes data for exports occurring between (at least some of the) country groups delimited by Kee, Nicita and Olarreaga. This data situation lends itself to the question of how exports have reacted to different levels of trade restrictions between groups of countries. However, the UNCTAD export data do not cover all groups of countries for which trade restriction information exists. The bold figures in the Table indicate those 20 trade relations for which we

have both types of information: on restrictions and on exports.

A reasonable first suggestion might be that export growth is negatively dependent on trade restrictions. For agriculture this relation should be expressed more clearly than for manufactured goods because the former are mainly standardized staple goods while the latter exhibit a greater degree of diversity and product differentiation, which might make trade restrictions less relevant for the dynamics of exports. The development of exports has been considered between 1993 and 2003 and is simply expressed as the 2003 value divided by the 1993 value. Trade restrictions are measured for the years between 2000 and 2004. It is assumed that their level does not change quickly and is, thus, relevant for the development of exports during 1993 to 2003.

Figure 1 depicts export dynamics and trade restrictions for agriculture goods. A certain negative relation is evident, as suggested. By contrast, Figure 2,

which is for manufactured goods, shows almost no relation, and if one is indicated, it has an implausible sign. If one takes agriculture and manufactured goods together (as in part A of the Table), the result (not shown as a graph) is likewise not clearly interpretable.

R. O.\*

## References

- Anderson, J. E. (2005), *Measuring the Restrictiveness of International Trade Policy*, MIT Press, Cambridge, MA.
- Busse, M. (2002), "Competition Intensity, Potential Competition and Transaction Cost Economics", *HWVA Discussion Paper* 183.
- Coppel, J. and M. Durand (1999), "Trends in Market Openness", *OECD Economics Department Working Paper* 221.
- IMF (2005), Review of the IMF's Trade Restrictiveness Index, February 14, 2005.
- Gehlhar, M. (1996), "Reconciling Bilateral Trade Data for Use in GTAP", *GTAP Technical Paper* 10.
- Kee, H. L., A. Nicita and M. Olarreaga (2004), *Ad-Valorem Equivalents of Non-Tariff Barriers*, Worldbank.
- Kee, H. L., A. Nicita and M. Olarreaga (2005), *Estimating Trade Restrictiveness Indices*, Worldbank.
- Laird, S. and A. Yeats (1988), "Trends in Non-Tariff Barriers of Developed Countries 1966–1986", *Worldbank Policy, Planning and Research Working Paper* 137.
- OECD (2004), *Analysis of Non-Tariff Measures: The Case of Prohibition and Quotas* (TD/TC/WP(2004)28/FINAL).
- OECD (1997a), "Measurement of Non-Tariff Barriers", *Economics Department Working Paper* 179-68.
- OECD (1997b), *Indicators of Tariff and Non-Tariff Barriers* (Update).
- UNCTAD (2005), *Methodologies, Classifications, Quantification and Development Impacts of Non-Tariff Barriers*, TD/B/COM.1/EM.27/2.
- UNCTAD, *Handbook of International Trade and Development Statistics*, various issues.

\* Carina Friedmann and Petya Stefanova, both graduate students of economics at the University of Nuremberg, interns at the Ifo Institute, did excellent data work for this article.



## NEW AT DICE DATABASE

In the fourth quarter of 2005 the DICE database ([www.cesifo.de/DICE](http://www.cesifo.de/DICE)) received entries in the following fields:

- Health Outcomes (e.g. absenteeism from work due to illness)
- Health Expenditure
- Taxing Foreign Trade.

The topic “Basic Country Characteristics” is now linked to EUROSTAT, where the most recent data on gross domestic product and its components are available.

The new topic “Experts’ Assessments” now comprises 83 tables.

## FORTHCOMING CONFERENCES

### International Institute of Public Finance (IIPF)

The 62nd Congress of the International Institute of Public Finance (IIPF) will be held at Paphos, Cyprus, 28–31 August 2006. The subject of the conference will be “Public Finance: Fifty Years of the Second Best – and beyond”. Prospective contributors are invited to submit papers or abstracts no later than 31 January 2006.

### International Society for New Institutional Economics

The 2006 Annual Conference of the International Society for New Institutional Economics (ISNIE) will take place at Boulder, Colorado, 21–24 September 2006.

### European Programme in Law & Economics

The European Programme in Law and Economics, an international graduate programme, is hosting a workshop on 10 February 2006, at the University of Hamburg.

### EALE Conference 2006

The 2006 meeting of the EALE is being organised by CERGE-EI, the Center for Economic Research in collaboration with the Institute of Economic Stud-

ies of the Faculty of Social Sciences at Charles University, the Czech Economic Society, the Czech National Bank and the EALE Secretariat in Maastricht. The conference will take place 21–23 September in Prague, Czech Republic. Papers are invited in any area of labour economics. Those who wish to present a paper are requested to send full (draft) papers to the EALE Secretariat, Maastricht, The Netherlands before 1 March 2006.

### European Society for Population Economics (ESPE)

The Twentieth Annual Conference will take place on 22–24 June 2006 at the University of Verona, Italy. The aim of the conference is to facilitate the exchange of research ideas and results across a range of fields, including the economics of private households, labour economics, public economics, demography and health economics. Barry Chiswick (University of Illinois at Chicago) will serve as the program chair, Federico Perali as the local organiser. All submissions should be sent by e-mail to the program committee, [espe2006@web.econ.uic.edu](mailto:espe2006@web.econ.uic.edu).

## WORLD VALUES SURVEY

The World Values Surveys were designed to provide a comprehensive measurement of all major areas of human concern, from religion to politics, to economic and social life, and two dimensions dominate the picture: (1) Traditional/secular-rational and (2) survival/self-expression values. The traditional/secular-rational values dimension reflects the contrast between societies in which religion is very important and those in which it is not. The second major dimension of cross-cultural variation is linked with the transition from industrial society to post-industrial societies, which brings a polarisation between survival and self-expression values.

Four waves of the Values Surveys were conducted, in 1981, 1990, 1995, and 1999–2001. The fifth wave of the World Values Survey went into the field on 1 July 2005 and will continue until late 2006.

**DICE**  
**Database for Institutional Comparisons in Europe**  
**[www.cesifo.de/DICE](http://www.cesifo.de/DICE)**

The database DICE was created to stimulate the political and academic discussion on institutional and economic policy reforms. For this purpose, DICE provides country-comparative information on institutions, regulations and the conduct of economic policy.

To date, the following main topics are covered: Labour Market, Public Finances, Social Policy, Pensions, Health, Business Environment, Natural Environment, Capital Market and Education. Recently a chapter on Experts' Assessments of Governance Characteristics has been added. Information about Basic Macro Indicators is provided for the convenience of the user.

The information of the database comes mainly in the form of tables – with countries as the first column – but DICE contains also several graphs and short reports. In most tables, all 25 EU and some important non-EU countries are covered.

DICE consists primarily of information which is – in principle – also available elsewhere but often not easily attainable. We provide a very convenient access for the user, the presentation is systematic and the main focus is truly on institutions, regulations and economic policy conduct. Some tables are based on empirical institutional research by Ifo and CESifo colleagues as well as the DICE staff.

DICE is a free access database.

Critical remarks and recommendations are always welcome.

Please address them to

[osterkamp@ifo.de](mailto:osterkamp@ifo.de)

or

[ochel@ifo.de](mailto:ochel@ifo.de)

